II·II

May, 1961

Industrial Management

REVIEW

School of Industrial Management

Massachusetts Institute of Technology

Industrial Management

REVIEW

of the School of Industrial Management

Massachusetts Institute of Technology

II·II

May, 1961

CONTENTS

Editorial. 3

Price Movements in Speculative Markets: Trends or Random Walks, by Sidney S. Alexander 7

Simulation and the Problem of Air Traffic Control, by Martin Greenberger. 27

The Place of a Business Game in the Marketing Curriculum, by William F. Massy, Peter S. King, and C. Davis Fogg 43

Management Development as a Process of Influence, by Edgar H. Schein 59

The *Industrial Management Review* is published twice a year by the Industrial Management Review Association, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge 39, Massachusetts.

Editors: Raymond E. Britt, Jr., Raghu Nath, and James A. F. Stoner.
Associate Editors: John V. Olszewski and Michael Roemer.

Subscriptions: $2.00 per year.
Reprints available upon request.

Manuscripts should be submitted to the Editors, *Industrial Management Review*.

Copyright 1961 by the Industrial Management Review Association.

Editorial . . .

. . . The business of business

"What is really at stake here," said Judge Ganey in sentencing seven electrical equipment executives to jail for antitrust violations, "is the survival of the kind of economy under which America has grown to greatness, the free enterprise system." He might have added that America's greatness is also at stake, for the logic of world politics has placed the burden of greatness upon the vitality of our economy. Since World War II the nation's leaders, of both parties, have built Western democracy's defense upon a policy of American strength and influence. In the areas of greatest concern, in Asia and Africa, the immediate need is far less ideological than material; so the economic aid of democratic nations speaks for democracy, and the economic vigor of the United States' economy becomes, among the uncommitted, the measure of comparative desirability between democracy and communism. As if to emphasize this point, Mr. Krushchev's latest manifesto speaks in terms of an economic struggle, and, with nuclear warfare a potential alternative, our leadership seems quite willing to join the battle on the field of economics.

However, democratic influence abroad becomes meaningless to us if democracy fails at home. One of the long-accepted tenets of American political philosophy is the dependence of democracy upon economic freedom, and a more recent interpretation emphasizes economic order and security. Enslavement to economic masters and fear of economic uncertainty become serious threats to the exercise of democratic freedoms and prerogatives. American democracy, then, has a dual dependence upon its economy; its continued influence depends on economic vitality; its continued existence depends on economic freedom and order. Our method of providing both economic strength and freedom has been the free enterprise system.

Grown to gigantic size and exercising a great amount of control over that system, the corporate sector of our economy has become one of the principal forces in the struggle to keep democracy healthy. Although the role of a defender was not sought, it has come as the inevitable consequence of growth and must be recognized as

fact. Shrinking from this responsibility will not alter its existence, and if, by ignoring it, the corporation misuses its evolved power, it must nonetheless share the blame for weakening democracy. Occupying a crucial and privileged place in our society, the private corporation may forfeit its autonomy if it fails its trust. And, simply by virtue of their position, corporate managers themselves have become trustees of American democracy.

For this reason, the proceedings in Judge Ganey's court were profoundly disturbing. Forty-four executives of some of the nation's most respected corporations were found guilty of violating the Sherman Act by conspiring to control a great portion of the electrical equipment market. The offenses were blatant and calculated; the law was precise and its intention clear: to protect the free enterprise system in which the convicted executives and their companies purportedly believe. But something far more subtle and sinister is involved. These men are not simply hypocritical, nor are they ignorant. Many of them were well aware of the trusteeship function inherent in their positions and they accepted the responsibility in private life. In their pathetic pleas for leniency, defense attorneys spoke repeatedly of church wardens, benefactors of crippled children and cancer victims, boys' club directors, and members of governors' committees. Yet one of these men was able to say, "No one attending the gatherings was so stupid he didn't know they were in violation of the law. But it is the only way a business can be run. It is free enterprise." Evidently something happens to a man when he steps from the responsibilities of private life into the protecting anonymity of corporate life. But are men who leave conscience at home a necessary consequence of the corporate form? There is considerable argument to maintain that they are.

The corporation is the American executive's environment and its existence is often seen quite apart from the men who compose it. In fact, the law institutionalized this separation by creating the legal fiction of a corporate entity with its limited liability. After years of struggle to rise in the corporation, the top executive is understandably committed to it. He has responded successfully to its system of rewards and punishment and, as promotion follows promotion, his original decision to identify his success with the corporation's rewards becomes reinforced. Rewards, of course, result from producing. In the case of a sales manager the index of success is sales dollars and, if volume is satisfactory or growing, the company will continue to reward. In the particular case of the electrical equipment sales managers, the methods used to maintain or increase sales were illegal, but the rewards kept coming. With no questions asked and within the sanctity of the corporation, the illegal mode of behavior was reinforced rather than discouraged.

In this, until recently, the legal system has been an accomplice. No law or legal code can function effectively without placing direct and inescapable responsibility upon individuals. Fines upon cor-

porations are easily absorbed and no individual need suffer; fines upon individuals can be passed on to their supporting corporations. However, a jail sentence provides that inescapable punishment upon which law depends. Until recently the Sherman Act provision for a jail sentence has not been used and the rewards of the corporation have outweighed the fears of punishment from conviction. Although the recent sentencings may help to emphasize the need for individual responsibility among executives, the basic problem is still not solved. Other decisions of corporate executives, affecting the welfare of many and, as argued before, the future of democracy, do not fall under criminal law. If the corporation has bred men who violate the criminal statute, might it not also tend to breed men who will act irresponsibly in other areas? Wherever the corporation's goals conflict with society's welfare (and this is certainly possible), its executives will be under pressure to act in an anti-societal manner in order to collect the immediate, tangible rewards of the corporation to which they are committed. In the framework of America's dependence upon its corporations for success in the world struggle, this behavior can be just as irresponsible as that of the law-breaking electrical executives.

The problem of the development of responsible managers will be insoluble as long as the corporation is seen as an impersonal institution which predetermines the behavior of the executives within it. The corporation is not, however, in any meaningful sense, an entity completely separate from the men who compose it. Some person or group must make the decisions that give a corporation its character, and someone must implement these decisions. While it is true that at certain levels and at certain times the complex of policies often seems like an untouchable, unalterable taskmaster, at some point human control does exist. When and where it does, individual responsibility should be clear. Judge Ganey recognized this fact by jailing only those men he thought were policy-makers. That these men were not mere pawns of policy but its determiners, is a greater indictment of their executive character and should therefore cause greater alarm. The question of the policy-makers above them has not been answered satisfactorily: either they, too, must have been guilty or their methods of control are woefully inadequate. But the policy-makers are not alone to blame: those who follow blindly and illegally must also share the guilt. The problem for the American corporation, in shouldering its responsibility to society, lies in its development of responsible individuals. Thoreau wrote, "It is truly enough said that a corporation has no conscience, but a corporation of conscientious men is a corporation with a conscience." But to whom does the conscientious executive owe his primary allegiance?

The key to an individual's responsibility lies in his frame of reference. The man whose frame of reference is limited to himself can be counted upon for anti-social behavior whenever his values conflict with society's. When he refers his values to a par-

ticular group, he may well become a partisan and is likely to act as the group expects. When he thinks in terms of society, he begins to build societal values into his motivation. Undoubtedly, any one man employs all three frames of reference in varying amounts. If, in its development of executives, the corporation rewards a man for emphasizing the corporation as his reference group to the virtual exclusion (or at least subjugation) of society as a frame of reference, irresponsible behavior of the type we have been discussing will be fostered. A sales manager who thought in terms of the public good in addition to his corporation's well-being would not have allowed himself to rig bids or otherwise to act illegally. More crucially, the corporation must not allow an executive to find, as Judge Ganey said, "balm for his conscience in additional comforts and the security of his place in the corporate setup." His responsibility to society is a personal one, and it is the corporation's job to encourage him to see it that way.

America's corporations cannot escape a great share of the responsibility for democracy's strength and, in fact, for the strength of Western civilization. Corporations are composed of men and, if they are to succeed, corporations must develop executives with a broad social outlook and a keenly-felt personal obligation, not only to their company but to society as well. The electrical equipment anti-trust convictions are appalling because they are symptomatic of the corporations' failure to meet this crucial need. With America's greatness at stake, profits alone cannot justify a corporation's existence. Times have changed and we can no longer afford to say, with Calvin Coolidge, that "the business of America is business." Today we must insist that "the business of business is America."

Price Movements in Speculative Markets: Trends or Random Walks*

by SIDNEY S. ALEXANDER, *Professor of Industrial Management*

There is a remarkable contradiction between the concepts of behavior of speculative prices held by professional stock market analysts on the one hand and by academic statisticians and economists on the other. The professional analysts operate in the belief that there exist certain trend generating facts, knowable today, that will guide a speculator to profit if only he can read them correctly. These facts are believed to generate trends rather than instantaneous jumps because most of those trading in speculative markets have imperfect knowledge of these facts, and the future trend of prices will result from a gradual spread of awareness of these facts throughout the market. Those who gain mastery of the critical information earlier than others will, accordingly, have an opportunity to profit from that early knowledge.

The two main schools of professional analysts, the "fundamentalists" and the "technicians," agree on this basic assumption. They differ only in the methods used to gain knowledge before others in the market. The fundamentalist seeks this early knowledge from study of the external factors that lie behind the price changes. In a commodity market he tries to estimate the future balance of supply and demand for that commodity. In the stock market he studies general business conditions and the profit prospects for various industries, and for the individual firms within those industries, with special attention to new developments.

The "technician" operates on the same basic assumption, that facts existing at one time will govern the prices at some future time, but he operates in a different manner. He leaves to others the study of the fundamental facts in the reliance that as those others act on their knowledge there will be a detectable effect on the price of the stock or commodity. The technician, accordingly, studies price movements of the immediate past for telltale indications of the movements of the immediate future.

Both schools of analysts thus assume the existence of trends which represent the gradual recognition by the market of emergent factual situations - trends which, if they exist, must depend for

* The data gathering and processing underlying this research were supported by the Sloan Research Fund of the School of Industrial Management, Massachusetts Institute of Technology.

their existence on a lagged response of the market prices to the underlying factors governing those prices. It might, at first blush, seem possible that the trends arise not from a lagged response of the market price to the fundamental circumstances, but rather from a trend in those underlying circumstances themselves. Thus, although a stock's price might at all times represent a given multiple of its earnings, its earnings might be subject to a long run trend. If, however, there are really trends in earnings, so that an increase in earnings this year implies a higher probability of an increase next year than do stable or declining earnings, the stock price right now should reflect these prospects by a higher price and by a higher ratio of price to current earnings. Consequently, if there is no lagged response there should be no trend in prices. By a trend in this connection we mean a positive serial correlation of successive changes or, more generally, a probability of future price change dependent on present price change.

The professional analysts would certainly not subscribe to the notion that the best picture of the future movements of prices can be gained by tossing a coin or a set of coins. Yet that is just what the academic students of speculative markets say is the best way. The academic students of speculative markets have come to deny the very existence of trends in speculative prices, claiming that where trends seem to be observable, they are merely interpretations, read in after the fact, of a process that really follows a random walk. A price can be said to follow a random walk if at any time the change to be expected can be represented by the result of tossing a coin, not necessarily a 50-50 coin, however. In particular, a random walk would imply that the next move of the speculative price is independent of all past moves or events.

This probabilistic view of speculative prices is consistent with the theoretical bent of economists who like to talk about perfect markets. If one were to start out with the assumption that a stock or commodity speculation is a "fair game" with equal expectation of gain or loss or, more accurately, with an expectation of zero gain, one would be well on the way to picturing the behavior of speculative prices as a random walk. But in fact, this picture of a speculative price movement is as much based on empirical findings as on theoretical predispositions. In a pioneer work Bachelier,[1] a student of the great French mathematician Poincaré, derived, in his doctoral thesis in 1900, a theory that speculative prices follow random walks, largely from the assumption of zero expectation of gain. He then compared the statistical distribution of price behavior expected from this theory with observed distributions of price changes of certain government securities (*rentes*) on the Paris *Bourse*, and he found a close correspondence between the observed distribution and that to be expected from his theory.

---

[1] M. L. Bachelier, Théorie de la Speculation, Gauthier-Villars, Paris, 1900.

The most impressive recent findings confirming the random walk hypothesis are those of Kendall.[2] He calculated the first twenty-nine lagged serial correlations of the first differences of twenty-two time series representing speculative prices. Nineteen of these were indexes of British industrial share prices on a weekly basis. (See Table 1). Two of the remaining three were cash wheat at Chicago, one weekly and one monthly, and the last was the spot cotton price at New York, monthly. Essentially, Kendall was asking with respect to each weekly series: How good is the best estimate we can make of next week's price change if we know this week's change and the changes of the past twenty-nine weeks and correspondingly for the monthly series?

Contrary to the general impression among traders and analysts that stock and commodity prices follow trends, Kendall found, with two or three exceptions, that knowledge of past price changes yields substantially no information about future price changes. More specifically, he found that each period's price change was not significantly correlated with the preceding period's price change nor with the price change of any earlier period, at least as far as he tested, up to twenty-nine periods. Essentially, the estimate of the next period's price change could have been drawn at random from a specified distribution with results as satisfactory as the best formula that could be fitted to past data. In the case of weekly wheat prices, that distribution was studied in detail and it turned out to be very close to a normal distribution.

There was one notable exception, however, to this pattern of random behavior of price changes. That was the monthly series on cotton prices in the United States since 1816 with, of course, a few interruptions for such events as the Civil War. For this series there did appear to be some predictability, and Kendall felt impelled to draw the moral that it is dangerous to generalize even from fairly extensive sets of data. For, from the behavior of wheat prices and the stock prices, one might have concluded that speculative markets do not generate autocorrelated price changes - and here was cotton providing a notable exception.

Alas, Kendall drew the wrong moral. The appropriate one is that if you find a single exception, look for an error. An error there was, for the cotton price series was different from the others investigated by Kendall. Almost all the others were series of observations of the price at a specified time - say, the closing price on Friday of each week. Each observation of the cotton series was an average of four or five weekly observations of the corresponding month. It turns out that even if the original data - the Friday closing prices - were a random walk, with successive first differences uncorrelated, the first differences of the monthly average of

---

[2] M. G. Kendall, "The Analysis of Economic Time Series - Part I: Prices," Journal of the Royal Statistical Society (Series A), Vol. 96 (1933), pp. 11-25.

TABLE 1

SERIAL CORRELATIONS OF FIRST DIFFERENCES OF
KENDALL'S STOCK PRICE INDEXES, 1928-1938,
FOR VARIOUS INTERVALS OF DIFFERENCING

Kendall's No.	Industry Group	First Order Serial Correlations. Differences Taken Over Interval of:					Second Order Serial Correlations
		1 Week	2 Weeks	4 Weeks	8 Weeks	16 Weeks	8 Weeks
1	Banks and Discount Companies	0.058	0.092	-0.067	-0.059	0.191	0.136
2	Insurance Companies	.052	.048	- .056	- .169	- .035	.049
3	Investment Trusts	.301	.450	.279	.020	.066	.064
4	Building Materials	.125	.044	- .087	- .075	.345	.271
5	Coal	.148	.156	- .014	- .186	- .074	.032
6	Cotton	.087	.179	.036	- .227	- .075	.123
7	Electric Light and Power	.181	.145	- .024	- .130	.219	.232
8	Gas	.096	.235	.103	.119	.360	.232
9	Iron and Steel	.088	- .005	- .028	- .122	- .009	.057
10	Oil	- .013	.027	.015	- .015	.073	.061
11	Total Industrial Productive	.195	.155	- .016	- .149	.055	.110
12	Home Rails	.010	- .030	- .002	.162	.072	- .019
13	Shipping	.053	.032	- .017	.111	.388	.325
14	Stores and Catering	.230	.145	- .082	- .218	.104	.199
15	Total Industrial Distributive	.237	.179	- .008	.033	.303	.241
16	Breweries and Distilleries	.034	.102	.046	.020	.152	.078
17	Miscellaneous	.200	.191	- .011	- .101	.267	.267
18	Total Industrial Miscellaneous	.177	.190	.033	- .069	.291	.252
19	Industrials (All Classes Combined)	.234	.207	.018	- .063	.279	.246

four or five of these weekly observations would exhibit first-order serial correlations of about the magnitude Kendall found for cotton.[3] So Kendall's exception vanishes, and we are left with the conclusion that at least for the series he investigated the serial correlations were not significantly different from zero.[4]

But the question immediately arises whether a week is not an inappropriate period of observation. The market analysts might protest that when they speak of a trend they are speaking of a smooth underlying movement on which is typically superimposed a lot of short-term fluctuation. With weekly observations the short-term fluctuations might very easily swamp the underlying trends. In particular, the give and take of the market leads to a phenomenon, recognized by all analysts, of reactions, usually called technical reactions, presumably associated with profit taking. These reactions are, of course, negatively correlated with the main price swings. That's what makes them reactions. Kendall's correlations, close to zero, could possibly be a consequence of the combination of the negative contributions of the reactions and the positive contributions of the trends.

The path of a speculative price might, accordingly, be represented by a sum of two components, a smooth underlying trend or cycle changing direction only infrequently, and a much shorter cycle of action and reaction. Under this hypothesis the first-order serial correlations of daily price changes might be negative, the first order correlation of weekly changes might be close to zero, while the first order serial correlations of monthly or bimonthly changes might be significantly larger than zero.

We can test this possibility by studying the first order serial correlations of Kendall's data using successively longer intervals of differencing. As we do so, and consider first the first order serial correlation of one week changes, then of two week, four week, eight week, and sixteen week changes, the influence of the reactions should become smaller and smaller and the trend effect, if there is one, should become dominant. The corresponding correlations, roughly calculated,[5] are given in Table 1.

---

[3] This point was independently discovered by the author and by Holbrook Working. The latter, however, had the pleasure of first publishing it in "Note on the Correlation of First Differences of Averages in a Random Chain," Econometrica, Vol. 28, No. 4, October 1960, pp. 916-918.

[4] Another possible exception may be noted for Kendall's Series 3, Investment Trusts, whose first five serial correlations were 0.301, 0.356, 0.158, 0.164 and 0.066. This series will be mentioned again below.

[5] Roughly, because they were computed, not from the original data, but from the serial correlations published by Kendall. Since successive serial correlations are based on fewer observations because of the necessity of sacrificing end terms, a certain "end term error" is introduced by this procedure.

While an occasional high value correlation occurs in Table 1 for intervals greater than one week, it must be remembered that for the given total time period under study, the number of observations drops in proportion to the length of the differencing interval. Since the variance of the correlation coefficient is inversely proportional to the number of observations, it is directly proportional to the length of the differencing interval. An occasional high computed value of the correlation becomes increasingly probable as the differencing period is lengthened, even if the true correlation is zero.

It must be concluded that the data of Table 1 do not give any substantial support to the hypothesis that, as differences are taken over longer and longer intervals, the first order serial correlations of the first differences generally increase.

Once again, Series 3, investment trusts, is an exception. It seems to have a particularly high level of serial correlation on a two-week period of differencing. Other occasional higher values among the various series, for changes at sixteen week intervals in particular, have to be seriously discounted although they suggest intriguing possibilities for further study.

These higher correlations for sixteen week changes (for Series 4, 8, 13, 15, and possibly 17, 18, and 19) proceed from a rather curious relationship that holds for 18 of the 19 series. The first order correlations on an eight-week basis tend to be algebraically smaller than the second order correlations (see Table 1). The implication of the existence of an eight week half cycle may be an interesting suggestion, although it could hardly be said to be established by the data.

One further attempt was made, in spite of Kendall's findings that the serial correlations were not significantly different from zero, to see if some nugget of systematic trend behavior might still be found in his data. It is possible that while the lagged autocorrelations of any series were not found to be significantly different from zero when taken one at a time, they might jointly form a pattern that is significant. A simple test of this possibility was attempted. A trend was fitted to the first differences of each stock price series by a Spencer 21-term moving average.[6] Then the ratio of the variance of the moving average to the variance of the first differences themselves was computed.

The variance ratios given in Table 2 are to be interpreted as follows. If each first difference lay exactly on the moving average

---

[6] See E. T. Whitaker and G. Robinson, The Calculus of Observations, (4th ed.) London 1944, p. 290, for the formula used. Actually, it was not necessary to fit the trends to the series themselves, but the variance of the moving average, expressed in units of the variance of the first differences, could be computed by applying the smoothing coefficients of the Spencer formula directly to the lagged serial correlations.

TABLE 2

RATIOS OF THE VARIANCE OF SMOOTHED FIRST DIFFERENCES TO THE VARIANCE OF UNSMOOTHED FIRST DIFFERENCES KENDALL'S STOCK PRICE INDEXES

(Smoothing Performed by 21 Term Spencer Moving Average)

Series	Industry Group	Ratios
1	Banks and Discount Companies	0.161
2	Insurance Companies	.158
3	Investment Trusts	.525
4	Building Materials	.162
5	Coal	.202
6	Cotton	.210
7	Electric Light and Power	.206
8	Gas	.230
9	Iron and Steel	.158
10	Oil	.145
11	Total Industrial Productive	.212
12	Home Rails	.138
13	Shipping	.151
14	Stores and Catering	.212
15	Total Industrial Distributive	.219
16	Breweries and Distilleries	.171
17	Miscellaneous	.221
18	Total Industrial Miscellaneous	.218
19	Industrials (All Classes Combined)	.232

trend line, that is, if the trend line were a perfect fit, the variance ratio would be unity. If, on the other hand, all the serial correlations of order greater than zero were identically zero, the expected values for a random walk, the variance ratio would be 0.143, the sum of the squares of the coefficients in the smoothing formula. It is, of course, possible for the ratio to be even less than 0.143.

Except for Series 3, the trend variance is not a much larger proportion of the original variance of the first differences than would be expected in the case of a random walk. It must be concluded that, with this exception, if trends exist in the first differences, they are very weak.

All in all, Kendall's data do seem to confirm the random walk hypothesis. Further work by Osborne[7] strengthens the random walk hypothesis from a different point of view. While Kendall

[7] M. F. M. Osborne, "Brownian Motion in the Stock Market," Operations Research, Vol. 7, No. 2, March-April 1959, pp. 145-173. See also comment and reply in Operations Research, Vol. 7, No. 6, November-December 1959, pp. 806-811.

worked with serial correlations for each series separately, Osborne worked with ensembles of price changes. Roughly stated, he found that the changes in the logarithms of stock prices over any period in a given market, principally the New York Stock Exchange, constituted an ensemble which appeared to be approximately normally distributed with a standard deviation proportional to the square root of the length of the period. This proportionality of the standard deviation of price differences to the square root of the differencing period is a characteristic of a random walk and had been pointed out much earlier by Bachelier. In Bachelier's case, however, the differences were arithmetic, while in Osborne's they were logarithmic.

It must be noted that Osborne's measurements do not concern trends in the prices of stocks but merely the statistical distribution of the changes in the logarithms, which, as Osborne pointed out, correspond quite closely to percentage changes. That they do not correspond exactly to percentage changes has an important bearing on one of Osborne's principal findings, as we shall see.

Osborne also supplied a theoretical mechanism that could explain the observed pattern of price movements. The mechanism is a random walk in the logarithm of prices with each step being a constant logarithmic value, depending on the time length of the step. The basic step is a transaction of which there might be ten or a hundred a day. The compounding of such steps in familiar probability sequences would, over any period of time, yield a normal distribution of changes in the logarithms of price, with standard deviation proportional to the square root of the period over which comparison is made.

One peculiar result of Osborne's proposed mechanism merits further study. Bachelier, the pioneer in regarding speculative price behavior as a random walk, derived the theoretical properties of the distribution of changes in the prices of *rentes* on the assumption of a "fair game," that is a zero expectation of gain. A price change in either direction of a given amount was equally probable in Bachelier's model. Osborne made a somewhat different assumption with a radically different result; he assumed that a change in either direction of a given amount in the logarithm of price was equally likely, no longer a fair game.

Thus, under Bachelier's assumption, given an initial investment value, say \$100, it would be equally probable, at the end of time T to be worth \$100+k or \$100 - k. Exactly how large k would be for any stated probability would depend on the fundamental constant of the distribution and the square root of the length of time T. But whatever the value of k, so long as the probability of a gain of k is equal to the probability of a loss of k for all k within the permitted range, the expected value in any future period remains \$100 and the expected gain 0.

Under Osborne's assumption, however, there will be an expectation of gain. Suppose, using logarithms to base 10 and starting from $100, a gain over some particular period of time, say five years, is equally likely to be a gain of logarithm 1 or loss of logarithm −1. These would correspond to an equal probability that the price at the end of five years would be either $10 or $1,000. The mathematical expectation in this case would be $505, or an expected gain of $405. This example illustrates the familiar difference between the arithmetic and the geometric mean. Over the long run, then, it makes a great deal of difference whether there is an expectation of zero arithmetic gain or zero logarithmic gain. In the latter case there will be a tendency for an investment value to grow, independent of any growth in the economy other than the growth implicit in the existence of a random walk in the logarithms.

How clearly established an empirical finding is the logarithmic rather than the arithmetic step in the random walk? Osborne was led to the logarithmic form, while Bachelier was not, because Bachelier considered only a single type of security at a time, whereas Osborne considered an ensemble of prices, usually all the common stock prices on a particular exchange. Osborne assumed, without much explicit consideration, that it was appropriate to try to fit one distribution of expected change to all common stocks, whether priced at $100 or $10 or whatever. Very little empirical investigation is required to show that the relative frequency of price rises of $10 in one month is much smaller among stocks selling at $10 than among stocks selling at $100. On the other hand, it is quite reasonable to expect that the relative frequency of a $1 price rise in a month among $10 stocks would be about equal to the relative frequency of a $10 price rise in a month among $100 stocks. And rough empirical tests seem to be in accord with the latter case.

If then we have to choose a single distribution that will fit stocks of all different prices, and if our only choice were between equal probabilities of dollar amount changes and equal probabilities of proportional changes, we are necessarily led to choose the latter. The assumption of equal probabilities of given changes in the logarithm of price falls in the latter class.

But there are other possible models which yield equal probabilities of changes of given proportions. One is of particular interest to us in that it certainly fits the data as well as the logarithmic model and does not imply a built-in growth of values as does the logarithmic. It postulates equal probability of given percentage changes, almost the same as equal probability of given logarithmic changes, but not quite. On the tiny difference hinges the existence or nonexistence of the remarkable property of speculation being a game biased in favor of winning.

In both the logarithmic form and the percentage form of the hypothesis it is equally probable that a $100 stock rises by $10 in

a month or that a $10 stock rises by $1. Under both schemes it is equally probable that a $100 stock declines by $10 in a month or a $10 stock by $1. But under the percentage form it is equally probable that a $100 stock goes to $101 or to $99 in a given time, whereas in the logarithmic form it is equally probable that a $100 stock goes to $101 or $99.01 in a given time. This difference of one cent in a dollar change from $100 spells the difference between zero expectation of gain and positive expectation.

If, then, the percentage hypothesis is adopted instead of the logarithmic hypothesis, the expectation of gain disappears. The difference between the distributions generated by the two hypotheses would, over most time periods of practical interest, be so fine that any test delicate enough to distinguish between them is likely to throw them both out.

In fact, both hypotheses would generate normal distributions of the changes in the logarithms of prices, differing only in their means and standard deviations. In testing various models of this sort we generally infer the mean and the standard deviation from the data and assume that the mean was influenced by general economic conditions separate from the random walk. Under these circumstances we can say whether the observed distribution is or is not close to normal, but we cannot say whether it is closer to the percentage hypothesis or to the logarithmic. To discriminate between these hypotheses we would need an independent measure of the random step, or of the standard deviation. Bachelier actually derived such an independent measure from the price of options, but Osborne merely showed that the distribution resembled a normal distribution and the standard deviation increased with the square root of the differencing period. To whatever extent his findings support the logarithmic hypothesis, they also support the percentage hypothesis.

But Osborne did not rigorously test the normality of the distribution. A rigorous test, for example the application of the chi-square test to some of the data used by Osborne, would lead us strongly to dismiss the hypothesis of normality. (See Table 3). It yields a chi-square of over 60 for 8 degrees of freedom, although almost all of the discrepancy between actual and expected frequencies arises from the extreme classes of increases or decreases greater than 10 per cent.[8] It may be presumed that special factors operated to produce far more large price changes than are characteristic of a normal distribution. This sort of situation (leptokurtosis) is frequently encountered in economic statistics and would certainly overshadow any attempt to test fine points such as the difference between a logarithmic and a percentage scheme.

---

[8] It should be noted that Osborne remarked that the tails of the observed distribution did not appear to correspond to those of the normal distribution.

TABLE 3

CHANGES IN PRICES OF LISTED COMMON STOCKS (NYSE)
OBSERVED COMPARED WITH EXPECTED

(Month Ending November 15, 1956)

Percentage Price Change	No. of Issues	
	Observed [a]	Expected [b]
+ 10% and over	54	21
+ 8% to + 10%	30	30
+ 6% to + 8%	50	54
+ 4% to + 6%	71	87
+ 2% to + 4%	119	127
- 2% to + 2%	346	328
- 4% to - 2%	149	157
- 6% to - 4%	100	121
- 8% to - 6%	74	79
- 10% to - 8%	31	43
- 10% or over	51	28
TOTAL	1075	1075

[a] Source: The Exchange (NYSE) Dec. 1956, back cover.

[b] On the assumption of normal distribution with mean equal to the median of the observed distribution, with standard deviation equal to the semi-interquartile range divided by 0.6745, and with total frequency equal to observed total frequency.

In any case the requirement that equal proportional gains for stocks of different prices should have equal probability does not imply a nonzero expectation. The nonzero expectation follows specifically from the assumption of unequal steps, measured in dollars, in the random walk, as between steps up and steps down. For small steps, equal logarithmic changes imply almost equal percentage changes, but the very small difference eventually grows to a very large one, another demonstration of the wonderful power of compound interest.

But more to the point than the difference between logarithmic and percentage schemes is the question: How far do Osborne's findings go to show that stock market prices really follow a random walk? Osborne, at best, merely showed that stock price changes might, to a rough approximation, have been generated by a random walk type of model. He suggests conditions that are sufficient to generate the observations, but not necessary. In plain language, he has a scheme which could have generated his observations; maybe it did and maybe it didn't.

It does seem that the principal clash between the analyst and the academic concerns not the nature of the statistical ensemble of price changes but rather the question of the existence of trends. For, if the implication of Kendall's findings are indeed general, it would make as much sense to try to predict the outcome of a coin toss as to predict movements of the stock market. The prime issue is, therefore, whether there is some way in which speculative price behavior is not random.

In order to attack this problem, the author ran off a set of simple tests of randomness of successive monthly or weekly changes of speculative prices. The technique may be illustrated by a weekly series on wheat (Wednesday closing prices of cash wheat over the period 1883-1934, excluding 1915-1920, a total of 2,379 weeks). Each week was classified as being a week of price rise or price fall, and the lengths of the runs were tallied. A run is defined as a sequence of successive weeks in which the price moves in the same direction. Table 4 shows the resulting table of runs and compares it with the distribution to be expected on the assumption of a random walk with equal probability of rise or fall.

TABLE 4

DISTRIBUTION OF LENGTHS OF RUN OF WEEKLY CASH WHEAT PRICES AT CHICAGO [a]

(1883-1934, Excluding 1915-1920)

Length of Run (Weeks)	Observed		Expected [b] Up or Down
	Up	Down	
1	280	295	297
2	147	132	149
3	86	77	74
4	38	42	37
5	15	18	19
6	13	12	9
7 or longer	7	8	9
TOTAL	586	584	595 [c]

[a] Source: Holbrook Working, "Prices of Cash Wheat and Futures at Chicago Since 1883," Wheat Studies Vol. II, No. 3, November 1934, pp. 75-124. See also Kendall, op. cit., Table 1 for frequency distribution for these prices, of differences between week t and t + 1 against differences between weeks t + 1 and t + 2.

[b] Expected on assumption of 0.5 probability of rise or fall, and 2,379 monthly observations.

[c] Expected total differs from sum because of rounding.

The correspondence is very close indeed, suggesting that at least the sequence of directions of changes in weekly wheat prices might have been produced by a random walk. Kendall[9] previously showed that, except for a few extreme items, the distribution of the size of these weekly wheat price changes is normal and independent of the movement of the preceding week.

The distribution of the runs of Standard and Poor's monthly composite index of stock prices,[10] 1918-56, is, however, inconsistent with the assumption of a random walk of equal probability of rise or fall. (Expectation 1 for 1918-56 in Table 5). We may note, however, that for this stock average for the period 1918-56:

(a) The relative frequencies of rising and declining months were 0.58 and 0.42, respectively;

(b) The relative frequency of rising months among all months for which the preceding month was rising, $p(+|+)$, was 0.67, and the relative frequency of declining months for which the preceding month was declining, $p(-|-)$, was 0.50.

A new set of expected runs based on these contingent relative frequencies for 1918-56 did fit the observed runs quite well. (Expectation 2 for 1918-56 in Table 5). Furthermore, expectations based on these relative frequencies derived from the 1918-56 data fit the 1871-1917 data for upruns quite closely, but downruns only fairly well. (Expectation 2 for 1871-1917 in Table 5). Actually, the simpler hypothesis (Expectation 3) that the probability of continuation of a run is 0.6 and of termination 0.4 fits the 1871-1917 data very well and also the 1918-56 upruns. It does imply a much higher frequency of long downruns than was observed in 1918-56, however.

Unfortunately, this evidence of the probability of one month's movement depending on the previous month's is entirely the result of using an average of weekly prices for each month's observation. Distributions of runs of industrial stock prices for 1897-1929 and 1929-59 are given in Table 6, based on a single-point observation for each month, the last trading day of the month. These distributions are very close to those to be expected on the assumption of each month's change being independent of the previous month's, with a 0.57 probability of a monthly rise and 0.43 of a decline. (Expectation 4, Table 6). The same probabilities seem to fit both periods equally well, and the hypothesis of trends seems to be blown sky high.

One notable exception to the rule that statisticians have not found trends in stock market prices is furnished by the work of

---

[9] Op. cit.

[10] The index is a monthly average of weekly observations.

TABLE 5

LENGTHS OF RUN OF STANDARD & POOR'S[a] MONTHLY COMPOSITE STOCK PRICE INDEX, OBSERVED (Obs) vs. EXPECTED (Ex)

(Monthly Averages of Weekly Indexes)

Length of Run (Months)	January 1918 to March 1956								January 1871 to December 1917					
	Up				Down				Up			Down		
	Obs	Ex 1[b]	Ex 2[c]	Ex 3[d]	Obs	Ex 1[b]	Ex 2[c]	Ex 3[d]	Obs	Ex 2[c]	Ex 3[d]	Obs	Ex 2[c]	Ex 3[d]
	(Number of Runs)													
1	31	58	30	37	39	58	45	37	41	37	45	40	56	45
2	21	29	20	22	27	29	23	22	29	25	27	30	28	27
3	16	15	13	13	15	15	11	13	15	17	16	17	14	16
4	10	7	9	8	9	7	6	8	11	11	10	9	7	10
5	5	4	6	5	1	4	3	5	4	7	6	4	4	6
6 or longer	9	4	12	8	1	4	3	8	9	15	8	10	4	8
TOTAL	92	117	90	93	92	117	90[e]	93	109	112	112	110	112[e]	112

[a] Source: 1918-1956, Standard & Poor's Trade & Securities Statistics (Security Price Index Record). Number of Monthly observations: 467 for 1918-56, and 559 for 1871-1917.

[b] Based on probability of 0.5 for rise or fall, successive months independent, derived from a priori hypothesis.

[c] Based on $p(+|+) = 0.67$ and $p(-|-) = 0.50$, derived from 1918-56 relative frequencies.

[d] Based on $p(+|+) = p(-|-) = 0.6$, derived from 1871-1917 relative frequencies.

[e] Expected total, differs from sum because of rounding.

TABLE 6

DISTRIBUTION OF LENGTHS OF RUN OF MONTHLY INDUSTRIAL STOCK PRICES,[a] OBSERVED (Obs) vs. EXPECTED (Ex)

(End of Month Closing Prices)

Length of Run (Months)	Jan. 1897 - Jan. 1929				Feb. 1929 - Dec. 1959			
	Up		Down		Up		Down	
	Obs	Ex 4[b]	Obs	Ex 4[b]	Obs	Ex 4[b]	Obs	Ex 4[b]
	(Number of Runs)							
1	42	40	58	53	38	38	49	50
2	24	23	18	23	22	22	19	22
3	16	13	14	10	9	12	10	9
4	6	7	4	4	7	7	4	4
5	1	4	1	2	3	4	2	2
6 or longer	7	6	1	1	8	6	2	1
TOTAL	96	93	96	93	87	88[c]	86	88

[a] Dow Jones Industrials, Jan. 1897 to Jan. 1929 with break from July 1914 to Jan. 1915 (379 months). Source: The Dow Jones Averages, Barron's, New York 1931; Standard & Poor's Industrials, Feb. 1929 - Dec. 1959 (362 months). Source: see fn. (a), Table 5.

[b] Based on assumption of independence of successive monthly changes, and on the probability of 0.57 of a rise, and 0.43 of a fall, derived from entire period 1897 - 1959.

[c] Expected total differs from sum because of rounding.

Cowles and Jones.[11] They found an excess of sequences over reversals in stock market prices, but they used, at least for monthly data, averages of weekly observations. Their findings presumably derive principally from this error, plus a second effect, the influence of more frequent movements up than down.

The first effect is illustrated by the data in Table 5, based on monthly averages of weekly observations. For these data there is a ratio of 1.56 of sequences to reversals for 1871-1917, and 1.54 for 1918-56. The ratio of sequences to reversals shrinks to 1.045 for the 1897-1959 end of month data of Table 6. The second effect can largely be explained on the basis of a 0.57 relative frequency of monthly rises and 0.43 of monthly falls for 1897-1959. For if each monthly movement were drawn independently at random with these probabilities of rise and fall, we should expect a ratio of

[11] Alfred Cowles and Herbert E. Jones, "Some a Posteriori Probabilities in Stock Market Action," Econometrica, Vol. 5, No. 280, July 1937, pp. 280-294.

sequences to alternations of 1.040, almost identical to that observed.

In a revision of his findings designed to eliminate the first effect, Cowles still found some evidence, though weaker, of the existence of an excess of sequences over reversals, a ratio of 1.07 for monthly one point observations as compared with an earlier figure of 1.33 for the monthly averages of weekly observations.[12] But he did not make any allowance for the effect of a higher relative frequency of rises than declines. If the relative frequency of monthly rises were between 0.59 and 0.60 in the period covered by Cowles' data, as is likely, the ratio of 1.07 would be expected even if monthly movements were independent.

It must be concluded that the month to month movement of stock prices, at least in direction, is consistent with the hypothesis of a random walk with about a 6 to 4 probability of a rise. Evidence to the contrary was spurious, arising from the correlations introduced by monthly averaging or neglect of the unequal probability of rise and fall.

Lest the reader take undue comfort from the bullish implication of the 6 to 4 probability of a rise, it must be noted that declines, though less frequent than rises, are sharper, so that the average decline is substantially larger than the average rise. Nevertheless, there has been a well recognized upward trend of stock market prices of the order of 5.6% a year over the period 1897-1959. (Table 7).

One final test may be reported that should give great comfort to the analyst and encouragement to those who would use statistical studies to guide their speculative efforts, for it furnishes evidence that stock price changes could not have been generated by a random walk. Suppose we tentatively assume the existence of trends in stock market prices but believe them to be masked by the jiggling of the market. We might filter out all movements smaller than a specified size and examine the remaining movements. The most vivid way to illustrate the operation of the filter is to translate it into a rule of speculative market action. Thus, corresponding to a 5% filter we might have the rule: if the market moves up 5% go long and stay long until it moves down 5% at which time sell and go short until it again moves up 5%. Ignore moves of less than 5%. The more stringent the filter, the fewer losses are made, but also the smaller the gain from any move that exceeds the filter size. Thus with a 5% filter there will be a loss on any move between 5% and 10.53% and a gain on any move larger than 10.53%. For if the move is just a 10.53% move, say from 100 to 110.53, then we would

---

[12] Alfred Cowles, "A Revision of Previous Conclusions Regarding Stock Price Behavior," Econometrica, Vol. 28, No. 4, October 1960, pp. 909-915. The data cover about 1,000 months over the periods 1834-1865, 1897-1922, 1928-1958.

go long at 105 (100 plus 5%) and sell at 105 (110.53 minus 5%) and so just break even. With a 10% filter most of the moves which entailed a loss with the 5% filter would be filtered out. But a 20% move, which would yield a 9% profit with a 5% filter (computed on the lower vertex of the move, actually about 8.6% of the purchase price), would yield a 2% loss on a 10% filter.

Thus, as the filter size is increased, the number of transactions is reduced, and losses on small moves are eliminated, gains on large moves are reduced, and some moves which would yield gains with a small filter will yield losses with a large. This example illustrates the familiar tradeoff between reliability of the information and the cost of the information. The more stringent the filter, the higher the reliability, but the more of the move that is sacrificed in identifying it both in getting in and in getting out.

The results of the application of various filters to the Dow Jones and Standard & Poor's industrial averages from 1897 through 1959 are shown in Table 7. If stock price movements were generated by a trendless random walk, these filters could be expected to yield zero profits, or to vary from zero profits, both positively and negatively, in a random manner. Given an underlying long term trend, they might be expected to produce some profits, the greater profits being associated with the greater filter, but in any case, profits smaller than could be expected from just buying and holding.

In fact, medium filters uniformly yield profits, and the smallest filters yield the highest profits, and very high they are.

The retrospective gains from the filter rule (before commission) are compared in Table 7 with the gains that could be achieved over the corresponding period by just buying and holding. The results uniformly favor the smaller filters over the buy and hold method. Thus, the filter method derives its success from a characteristic of stock price behavior other than that implied by the upward long term trend alone. This conclusion is also confirmed by the fact, apparent from inspection of the work sheets, that the filter method made gains on the declines as well as on the rises.

From a practical standpoint these profits would be substantially reduced, but by no means eliminated, by the payment of commissions. I leave to the interested reader the computation of allowance for commissions.

It must be concluded that there are trends in stock market prices, once the "move" is taken as the unit under study rather than the week or the month. That is, the nonrandom nature of stock price movements revealed by Table 7 proceeds not only from filtering out small moves, but also from transforming the measure over which changes are considered. The many statistical studies which have found speculative prices to resemble a random walk have dealt with changes over uniform periods of time. The filter operation, however, deals with changes of given magnitude irrespective of the

TABLE 7

PROFITS FROM FILTERS OF VARIOUS SIZES COMPARED WITH BUY AND HOLD [a]

(1897 - 1959) [b]

Period	Filter Size [c] 5%	6%	8%	10%	12.5%	15%	20%	25%	30%	40%	50%	Buy & Hold
					Average Move (%) [d]							
1897 - 1914	13.8	15.8	19.8	22.8	30.7	39.6	62.6	62.6	82.5	80.2	97.0	-
1914 - 1929	12.8	14.9	19.7	25.4	33.3	43.0	69.4	115.8	115.8	115.8	115.8	-
1929 - 1959	14.5	16.4	22.3	26.3	31.6	36.1	52.9	72.3	188.9	199.0	291.0	-
					Average Profit Per Transaction (%) [e] (Before Commissions)							
1897 - 1914	2.9	3.0	2.7	1.5	3.2	5.4	12.2	4.0	7.7	(9.2)	(15.5)	75.3
1914 - 1929	2.0	2.2	2.6	3.6	5.2	7.8	16.3	32.5	24.7	9.6	(5.7)	596.6
1929 - 1959	3.5	3.6	4.8	4.3	3.9	2.9	6.0	9.8	11.2	43.2	57.3	154.1
					Number of Transactions [f]							
1897 - 1914	117	95	67	53	32	22	12	12	8	7	7	1
1914 - 1929	112	93	59	40	28	19	10	6	6	6	6	1
1929 - 1959	274	228	144	113	86	70	40	26	20	8	6	1
					Average Transactions Per Year							
1897 - 1914	6.5	5.4	3.8	3.0	1.8	1.2	0.7	0.7	0.4	0.4	0.3	-
1914 - 1929	6.6	6.3	4.0	2.7	1.9	1.3	0.7	0.4	0.4	0.4	0.4	-
1929 - 1959	9.0	7.5	4.7	3.7	2.8	2.3	1.3	0.8	0.6	0.3	0.2	-
					Average Profit Per Year (%) [g] (Before Commissions)							
1897 - 1914	20.5	17.4	10.5	4.6	5.8	6.6	7.8	2.6	3.2	(3.3)	(3.9)	3.2
1914 - 1929	15.8	14.7	10.7	10.0	9.9	9.9	10.3	11.1	8.6	3.4	(2.1)	14.1
1929 - 1959	36.8	30.0	24.5	16.8	11.4	6.9	7.8	8.2	7.0	9.3	8.5	3.0

(Parentheses Signify Losses)

(Parentheses Signify Losses)

[a] Based on Dow Jones Industrials, 1897 - 1929 and Standard & Poor's Industrials, 1929 - 1959. See fn. (a), Table 5, and fn. (a), Table 6 for source references.

[b] Periods: January 2, 1897 to July 30, 1914
December 12, 1914 to September 3, 1929
September 7, 1929 to December 31, 1959.

[c] 5% filter here designates 5% in either direction, others designate indicated percentages upward and equal logarithmic moves downward. E.g., 10% filter implies 10% upward or 9.09% downward.

[d] Calculated as follows: for each move, as defined by a specified filter, let the variable, M, denote the difference between the logarithms of the upper and lower endpoints of the move. The figure given as the average move is 100 (antilog $\overline{M}$-1) where $\overline{M}$ is the arithmetic mean of the M's.

[e] Let $\overline{R}$ be the average logarithmic profit defined as $\overline{R} = \overline{M} - 2F$. $\overline{M}$ is defined in fn. (d) and F is log (1 + f) where f is the filter expressed as a ratio, e.g., f is 0.10 for a 10% filter. The quantity 2F corresponds to the portion of the move that is used up in getting in or out. On an upmove of average size $\overline{M}$ the percentage profit would be $P_u = 100(\text{antilog } \overline{R}-1)$; on a downmove, $P_d = 100\, P_u / (100 + P_u)$.
The average profit entered in Table 7 is $\overline{P} = (100 + P_u)^{\frac{1}{2}}(100 + P_d)^{\frac{1}{2}} - 100$.

[f] A transaction is defined as a purchase and sale, so that each transaction would require two commissions. In each period there is one terminal transaction, such as for Dec. 31, 1959, terminated not by a filter signal but by the period limits. The corresponding terminal move was counted as half a move in the computation of $\overline{M}$, and of the number of transactions per year.

[g] Computed as $100\left[1 + (\overline{P}/100)\right]^{\overline{q}} - 100$, where $\overline{q}$ is the average number of transactions per year, and $\overline{P}$ is defined in fn. (e).

length of time involved. In short, it substitutes the dimension of the "move" for the dimension of time.

The findings surveyed in this paper can be summarized by the statement that in speculative markets price changes appear to follow a random walk over time, but a move, once initiated, tends to persist. In particular, if the stock market has moved up x per cent it is likely to move up more than x per cent further before it moves down by x per cent. This proposition seems to be valid for x ranging from 5 per cent through 30 per cent. It will require further study to find out if it is valid for x smaller than 5%.

The riddle has been resolved. The statisticians' findings of a random walk over the time dimension is quite consistent with non-random trends in the move dimension. Such a trend does exist.

I leave to the speculation of others the question of what would happen to the effectiveness of the filter technique if everybody believed in it and operated accordingly.

Simulation and the Problem of Air Traffic Control

by MARTIN GREENBERGER, *Assistant Professor of Industrial Management*

The collision of a United DC-8 jet with a TWA Super Constellation over New York harbor last December 16 dramatized for the public a question which has been troubling aviation experts for a number of years. Are present control measures realistic in view of the much greater speeds, much heavier demands, and much larger volumes brought about by jet traffic?

The December 16 crash took the lives of 134 persons, a number of fatalities without precedent in aviation history. But even more alarming and ominous than the severe death toll was the fact that both planes were presumably flying under "air traffic control" at the time of collision. That is, each plane was operating under "instrument flight rules" (IFR), and was therefore receiving flight instructions by radio from controllers on the ground throughout the course of its travel.

As far as authorities can recall, in every previous mid-air crash at least one of the planes involved had been flying under "visual flight rules" (VFR) on the principle of "see-and-be-seen." The blame has generally been attributed to this fact, causing a great deal of controversy about whether planes operating under VFR should be allowed access to controlled airways. The small-plane interests have lobbied hard for retention of mixed IFR-VFR traffic, while more progressive elements have recommended a division of the airspace into clearly defined zones which enforce a strict separation of visual from instrument flights.

In any case, as a result of the December 16 (all IFR) crash, the public's concern is currently diverted from the problem of VFR traffic to the more central issue of the adequacy and effectiveness of the entire air traffic control system as it now exists. This is certainly not the first time that this issue has received widespread attention. It was about five years ago that former President Eisenhower appointed Edward P. Curtis as his Special Assistant for aviation facilities planning, to undertake "the direction and coordination of a long-range study of the Nation's requirements for aviation facilities." In his charge to Mr. Curtis on February 10, 1956, Mr. Eisenhower wrote,

"I am taking this action because the rapid technical advances in aviation and the remarkable growth in the use of air transportation have confronted the Nation with serious aviation facilities problems. Modern aircraft can be operated in the numbers required by the national defense and the civilian economy only if airports, navigation aids, air traffic control devices and communications systems are suitable for their needs. Moreover, anticipated further increases in air traffic, the introduction of jet propulsion for civil as well as military aircraft, the advances being made in vertical flight, and the greater use of higher altitudes, all presage much heavier future demands upon our facilities for navigation and traffic control."

These words were written almost three years before the first jet was introduced to domestic passenger service by National Airlines (between New York and Miami) on December 10, 1958. As the words clearly indicate, our government was already fully aware in 1956 of the tremendous problems that were to accompany the use of jets commercially. A major effort was initiated to study these problems and make recommendations for their solution. Some of the ablest men in the country were called in to take part in this effort. The result of their deliberations was the "Curtis Report," presented to the President on May 10, 1957.

An introduction to the report contains the warning that

"Between the combined demands of civil and military aviation the existing control apparatus is already taxed to saturation. ... At the present time, our aviation facilities are hard-pressed to manage, within the bounds of prudence, even the fair-weather flow of traffic now contending for room on the principal routes and at the metropolitan approaches. They are wholly incapable, in their present form, of meeting more than a fraction of the far heavier demands upon our airspace which the national strategy and the market are certain to impose by 1975."

The warning was based on a forecasted growth in the total civilian aircraft population of from about 90,000 planes in 1957 to 125,000 planes in 1975. In view of the actual experience in the four years since the report was prepared, the warning seems mild, if anything. In these four years an increase of about 20,000 planes had already occurred, and the commercial accident rate (passenger deaths per hundred million miles flown) has almost trebled.

FEDERAL AVIATION AGENCY

The Curtis Report made several specific recommendations for government action. The two most important were to create:

1. immediately, but on an interim basis only, an Airways Modernization Board "responsible for developing communications, navigational aids, and control needed to accommodate future air traffic."

2. later, and on a permanent basis, an independent Federal Aviation Agency "into which are consolidated all the essential management functions necessary to support the common needs of the military and civil aviation of the United States."

Both of these recommendations were adopted. When the Federal Aviation Agency (FAA) was established in 1958, it completely absorbed the functions of the Airways Modernization Board as well as the former Civil Aeronautics Administration. A retired air force general, Elwood R. ("Pete") Quesada, was placed at the helm and expansion was undertaken. Today the FAA numbers just under 40,000 men.

When Quesada first assumed his responsibilities as FAA chief, he optimistically predicted installation by 1960 of a computer-radar system for the New York area which would relieve the controller's job of much of its strain and tedium. New York is the largest airline passenger generating area in the world. At peak periods it may have on the order of 200 planes simultaneously airborne, and this number is growing all the time. There is little doubt that the present manual system of control (designed 24 years ago when only 8 men were required to regulate the entire flow of air traffic across the country) is totally unmatched to the needs. The December 16 disaster, taken together with the alarming number of "near-misses" reported daily, bears startling testimony to this fact.

Quesada revised his prediction just before he relinquished his FAA post to become president of the Washington Senators baseball team and a director of American Airlines. He conceded that it could be another ten years before the automatic equipment now being engineered was fully operational.

The FAA checks out its new equipment and procedures at a 6000-acre experimental center set up for the purpose near Atlantic City. Present research and development effort includes work on three-dimensional radar and on anti-collision devices to help avert accidents. It also includes construction of electronic computing and communication systems to relieve the controller of such routine duties as keeping track of flights in progress and providing these flights with clearance information. The unexpectedly long time it is taking to put these systems into service is attributable largely to technological lead time; at least that is the most obvious reason. But there is also a second source of delay, this one inertial in nature, which will begin to become more noticeable once the systems are perfected and proved. As a matter of fact, the total delay due to inertia could eventually exceed the technological lead time.

Improvements in air traffic control up to the present time have always been within the framework of the old, manual system. "A body at rest tends to remain at rest, and a body in uniform motion tends to remain in uniform motion in the same straight line." But what is needed to alleviate the critical air traffic problem is an accelerated effort along a very different line from the one which has been followed traditionally.

Let us assume that the increased automation provided by the new equipment, when ready, is a first step in the correct direction. Then what is still needed is to adapt the present control system to

the new equipment (and not the other way around!) so as to ensure that the system will function in the most effective and safe manner possible. What we must do is to:

1. review methodically the whole system of control procedures, separation standards, airway configurations, altitude allocations, and terminal operations, to find possibilities for major improvement;

2. examine carefully proposed changes to determine whether they are feasible, safe, and indeed improvements;

3. undertake a comprehensive educational program to train air traffic controllers "before the fact" in the use of the new equipment and familiarize them with the changes decided upon.

Lofty aims like these require a research effort of major proportions, as the FAA is well aware. One of the chief needs is for a way of "dry running" a new procedure or piece of equipment without having to assume the obvious risks which a real trial implies. The approach most commonly taken to meet this need is to build an operable model or simulation of the real system in a form which permits the desired degree of experimentation. The simulation may then be used to review and examine; it may also be used to generate experience for teaching purposes.

SIMULATION

Simulation, as a general approach, is anything but new in the field of aviation. Flight simulators have been used for many years to test pilot trainees on the ground and to prepare them for their first experience in the air. Wind tunnels are a second familiar method of simulating flying conditions; they are generally used to inspect the design of airplane structures. There are numerous other aeronautical examples of simulation which could be cited.

But there is one form of simulation which was relatively unknown until about 1955. This is the simulation which uses a digital computer as a central component and represents the system or subsystem being simulated by a computer program. Besides the computer, other components of this type of simulation could include analog devices and people. For instance, the computer might represent a fabricated air traffic situation across the United States, while real controllers with telephones and with radar scopes fed by the simulation enact their customary (or deliberately altered) roles. Pilots in flight simulators could be added to the mock system, and so could further computers (or, alternatively, other programs on the same computer) for record keeping and automatic feedback of clearance data.

A simulation such as this, needless to say, would require an enormous amount of preparation and planning. Nothing of this scope has been attempted to date, but consideration has been given to it and there have been limited simulations at the FAA's experimental center near Atlantic City. One discussion of some of the

possibilities is contained in a Master's thesis by Leon Ephross, a former student at the MIT School of Industrial Management.

The aspect of an air traffic simulation which is sometimes the most perplexing to the uninitiated is at the same time one of the simplest to program for a computer. This is the mechanism which simulates the flow of aircraft into and out of a terminal, or through an en-route point between terminals. The flow mechanism must take account of the variations in average traffic densities both geographically and temporally. It must also provide for the uncertainties normally found in spacings between aircraft arrivals. It is an everyday occurrence, for example, that 3 planes arrive at an airport almost simultaneously, even though the mean arrival rate, when averaged over several hours, is 6 planes per hour; or that no planes arrive in a full 30 minute period.

In the process of simulating the arrivals and departures of planes at terminals, as well as the en-route flights of planes between terminals, a simulation makes Monte Carlo samplings from certain appropriate probability distributions to represent the uncertainties of the real situation. The remaining discussion treats in an elementary manner, first, the general concept of Monte Carlo sampling, and second, specific distributions which might be sampled during simulation.

MONTE CARLO SAMPLING

The behavior of certain physical processes, such as the occurrences in time of airplane arrivals at an airport, is often capable of representation by one or more probability distributions. The word stochastic is frequently used to describe this type of process when its behavior does admit of a probabilistic interpretation or representation.

One of the simplest stochastic processes to imagine is one in which arrivals (let us say) occur with time gaps of anywhere from 0 to 1 (hours, for example), and there is no greater likelihood of one size gap than of another within this range. The probability distribution which portrays this process is the uniform or rectangular distribution in the unit interval (0, 1). The density function* of the uniform distribution, f(u), is shown in Figure 1.

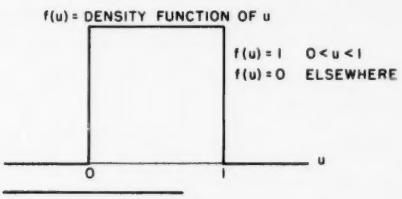

Figure 1. The uniform or rectangular distribution in the unit interval (0, 1).

* The density function of a given distribution, when integrated over an interval, expresses the likelihood that a sampling made at random from the given distribution will fall within the interval of integration. The function takes its name from the density function which depicts the distribution of mass along a linear rod.

In order to simulate arrivals under the aforementioned process, it suffices to make successive random samplings from the uniform distribution described by f(u). Suppose that the initial arrival of the simulation is known to occur at 1:00 p.m., and suppose that the first five samplings are 0.27, 0.71, 0.69, 0.03, and 0.58. The next five arrivals then would be set at,

$$1{:}00 + 0.27(60) = 1{:}16 \text{ p.m.}$$
$$1{:}16 + 0.71(60) = 1{:}59 \text{ p.m.}$$
$$1{:}59 + 0.69(60) = 2{:}40 \text{ p.m.}$$
$$2{:}40 + 0.03(60) = 2{:}42 \text{ p.m.}$$
$$2{:}42 + 0.58(60) = 3{:}17 \text{ p.m.}$$

It is assumed that successive gap sizes are completely independent.

This procedure of random sampling to form a schedule of arrivals is one of the most straightforward examples of a Monte Carlo simulation. How are the random samples taken? For the uniform distribution, they are taken merely by consulting one of the numerous tables of random numbers (if the simulation is conducted manually), or by a simple, programmed rule (if the simulation is conducted on a digital computer). The uniform distribution, however, is not generally a realistic representation of an arrival process. For the more realistic and more complicated distributions, samples may be taken by an application of a fundamental principle of the Monte Carlo technique. This principle provides a way of converting a uniformly distributed random number, obtained from a table or programmed rule, into a random sampling from an arbitrary probability distribution.

THE FUNDAMENTAL PRINCIPLE

Suppose that a random sampling is desired from a given probability distribution with a density function g(x) and cumulative function G(x)*. An inverse cumulative function $G^{-1}(y)$ may be defined by the condition that $x = G^{-1}(y)$ if and only if, $y = G(x)$. The fundamental principle states that if u is a random sampling from the uniform distribution in the unit interval, then $x = G^{-1}(u)$ is a random sampling from the distribution with cumulative function G(x). This principle holds whether the latter distribution is discrete or continuous, and whether it is empirical (measured) or analytical (theoretical).

---

* These two functions are connected by the relationship,

$$G(x) = \int_{-\infty}^{x} g(z)dz.$$

The cumulative function expresses the likelihood that a sampling made at random from the given distribution is less than or equal to the argument x. The cumulative function is non-decreasing as x increases, with values rising monotonically from 0 to 1. A unique, single-valued inverse function therefore exists.

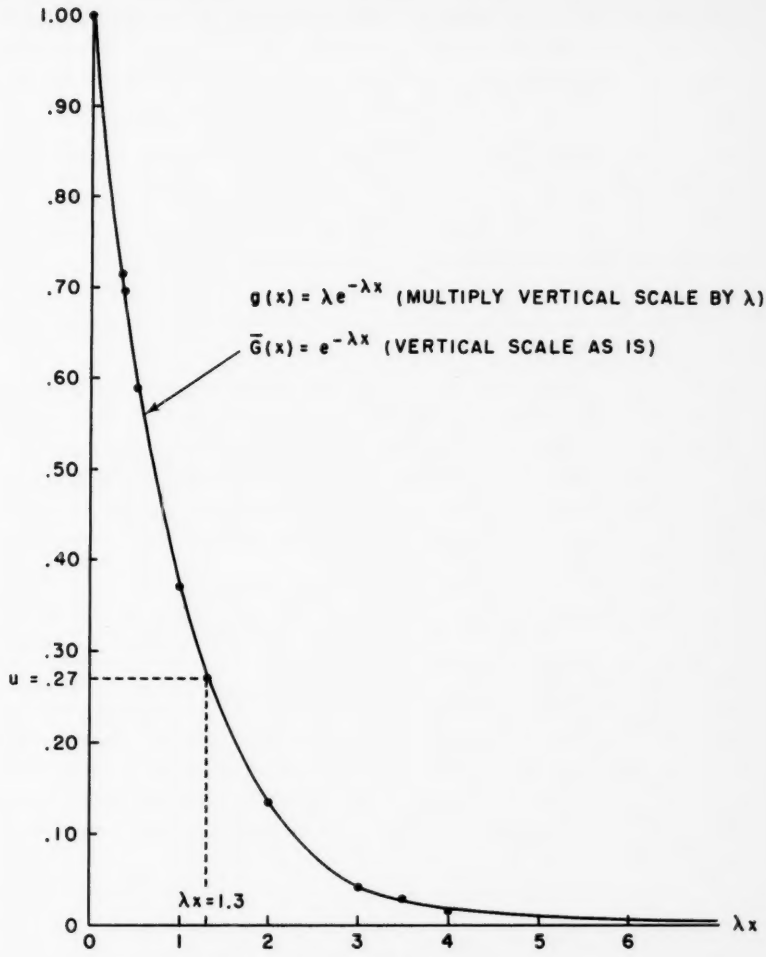

Figure 2. The standard exponential curve.

To illustrate the principle, consider the exponential distribution shown in Figure 2. The exponential distribution provides an exact representation of inter-arrival gaps when arrival times are "completely random;" that is, independent of time as well as of each other. The density function of the exponential distribution may be written,

$$g(x) = \lambda e^{-\lambda x}, \tag{1}$$

where the arrival rate λ is the single parameter of the distribution, and $1/\lambda$ is the mean of the distribution. The corresponding cumulative function may be written,

$$G(x) = 1 - e^{-\lambda x}. \tag{2}$$

When working with the exponential and related distributions, it is convenient to introduce a third function, which may be called the complementary cumulative function*, defined in general by,

$$\bar{G}(x) = 1 - G(x) \tag{3}$$

and given for the exponential distribution, in particular, by

$$\bar{G}(x) = e^{-\lambda x}. \tag{4}$$

If, when applying the fundamental principle, $G(x)$ is replaced by $\bar{G}(x)$, the validity of the principle remains intact. Because of the simplicity of its form, $\bar{G}(x)$ is used instead of $G(x)$ in our applications of the fundamental principle to the exponential and related distributions.

In order to draw a random sampling x from the exponential distribution, a random sampling u first is drawn from the uniform distribution in the unit interval. By equation (4), it is seen that in the case of the exponential distribution we must set,

$$u = \bar{G}(x) = e^{-\lambda x} \tag{5}$$

and therefore,

$$x = \bar{G}^{-1}(u) = \lambda^{-1} \log_e (u^{-1}) \tag{6}$$

Equation (6) provides one way to evaluate $x = \bar{G}^{-1}(u)$ for the special case of an exponential distribution. Another way, which works for any distribution, is by locating u on the vertical axis of a graph of $\bar{G}(x)$ (See Figure 2), then reading across horizontally until reaching the curve, and finally reading down vertically to the horizontal axis to find the value of x corresponding to u. This graphical technique is equally simple whether performed manually or programmed as a table look-up (and interpolation) on a digital computer.

The technique is carried out in Figure 2 with $u = 0.27$ as an example, giving $\lambda x = 1.3$ and $x = 1.3/\lambda$. This result holds for any λ, pointing up the fact that the curve in Figure 2 is an expression of any exponential distribution, regardless of the value of its parameter λ. The curve of Figure 2 is called the standard exponential curve and, with appropriate scaling of the vertical axis, may be interpreted as either the density function or the complementary cumulative function of any exponential distribution.

Suppose it now is desired to simulate a completely random arrival process whose mean gap size is 1/2 hour (i.e., $\lambda = 2$). As before, assume that the initial arrival of the simulation is known to occur at 1:00 p.m. and that the first five samplings from the uniform distribution are 0.27, 0.71, 0.69, 0.03, and 0.58. Employing

---

* The complementary cumulative function expresses the likelihood that a sampling made at random from the given distribution is greater than the argument x. It is non-increasing as x increases, with function values falling monotonically from 1 to 0.

Figure 2 and the graphical technique yields five random exponential drawings of 0.65, 0.17, 0.19, 1.75, and 0.27, and therefore successive arrivals of,

$$1{:}00 + 0.65(60) = 1{:}39 \text{ p.m.}$$
$$1{:}39 + 0.17(60) = 1{:}49 \text{ p.m.}$$
$$1{:}49 + 0.19(60) = 2{:}00 \text{ p.m.}$$
$$2{:}00 + 1.75(60) = 3{:}45 \text{ p.m.}$$
$$3{:}45 + 0.27(60) = 4{:}01 \text{ p.m.}$$

It is of some interest to observe the differences between these times and those of the earlier arrival process, noting that in both cases the theoretical mean gap size is 1/2 hour; that is, in both cases the theoretical arrival rate is 2 planes per hour.

DISTRIBUTIONS USED IN SIMULATING AIR TRAFFIC AT THE TERMINALS

The question of how to sample randomly from an arbitrary probability distribution has been answered. We consider next the question of which probability distributions in particular may be used in simulating air traffic, first in the terminal areas, and then in the en-route areas between terminals. Each distribution will be specified by its complementary cumulative function, $\bar{G}(x)$, which as we have seen is all that is required to apply our sampling procedure.

In a terminal area, estimated times of arrival (ETA's) are assumed, to a first approximation, to occur completely at random. This assumption is supported by several data studies which have been made in recent years. One way to improve the assumption is to remove the point of arrival, by convention, from the terminal itself to a periphery several miles outside of the terminal. It seems reasonable that the more distant this periphery is from the terminal, the more completely random are the ETA's. However, the fact that ETA's are estimated times of arrival, which for the most part are established prior to the exercise of scheduling measures by the terminal, argues in defense of the assumption, even for a periphery placed as close as the physical location where an aircraft begins its final landing procedure. For our purposes, therefore, an ETA at a terminal will mean the time at which an arriving aircraft would initiate its final landing procedure were it not for control imposed by the terminal for scheduling reasons.

Estimated times of departure (ETD's) from a terminal also are required in the simulation. ETD's refer to the time at which a departing aircraft would begin its take-off procedure were it not for other aircraft in and around the terminal area. ETD's may be obtained either from a given schedule of requested departure times or, in case the randomness assumption applies, from a Monte Carlo sampling identical to that used for ETA's. Needless to say, there are many other possible models which might be used to obtain ETA's and ETD's, but there is not sufficient evidence at present to warrant their inclusion in the discussion.

Let us now consider the case of completely random ETA's. We already have noted that gaps between successive ETA's are in this case exponentially distributed, and the $\bar{G}(x)$ of Figure 2 may be used to sample for these gaps. If the last ETA is known to have occurred at 1:00 p.m., for example, and a specification of the next ETA is required by the simulation, a gap is sampled using Figure 2 and, when added to 1:00 p.m., provides the desired ETA.

Suppose the following, slightly different problem arises in the simulation. The given time is 1:00 p.m. (or any other arbitrary time), the time of occurrence of the last ETA (at or before 1:00 p.m.) is unknown, and a determination of the next ETA (after 1:00 p.m.) is required. The requirement here is to sample for a prolongation of a gap rather than for a gap itself. Since a prolongation is almost always less than its associated gap, and never greater than this gap, our intuition suggests that prolongations are consistently smaller than gaps and that the $\bar{G}(x)$ developed for sampling gaps will not serve for sampling prolongations.

Our intuition is misleading in this case, although it would be correct for any arrival process other than the completely random, exponential process which has been postulated. One important characteristic of the pure exponential process is that it has no "memory." At a given point in time, a knowledge of when the last ETA has occurred does not improve or in any way alter our knowledge of when the next ETA will occur. For this reason the prolongation of a gap at an arbitrary point in time may be sampled using the same $\bar{G}(x)$ of Figure 2 that was used in sampling for the gap itself.

A third problem with still another twist may help to put the first two problems in clearer perspective. Suppose that the time is currently 1:00 p.m. (or any other arbitrary time), the times of occurrence of the last ETA (at or before 1:00 p.m.) and the next ETA (after 1:00 p.m.) are unknown, and a sampling is required of the total gap between these two ETA's. Even though this sampling is for a gap, it should not be made from the exponential distribution in this instance. The reason lies in the fact that, since no ETA is known to occur at 1:00 p.m., the gap will tend to be larger than it would have been were 1:00 p.m. definitely an ETA. In fact, its expected value is twice as large.

This somewhat perplexing line of reasoning may be easier to accept after considering a simpler, but related and similarly deceptive, question. "Given that exactly one A train and one B train arrive at a station each hour, and given that the A train always precedes the B train by precisely five minutes, what is the probability at an arbitrary time that the next train to arrive will be a B train?"*

---

* The answer is $\frac{1}{12}$, not $\frac{1}{2}$, even though $\frac{1}{2}$ of the arrivals are B trains.

The gap between ETA's on either side of 1:00 p.m. may be viewed as the sum of two independently distributed prolongations: a "backward prolongation" from 1:00 p.m. to the last ETA, and a "forward prolongation" from 1:00 p.m. to the next ETA. Each of the prolongations is exponentially distributed. Their sum is not, however, as already has been noted. It is simple to show mathematically that the sum has a distribution with density function,

$$g(x) = \lambda^2 x e^{-\lambda x}, \tag{7}$$

and complementary cumulative function,

$$\bar{G}(x) = (1 + \lambda x)e^{-\lambda x}. \tag{8}$$

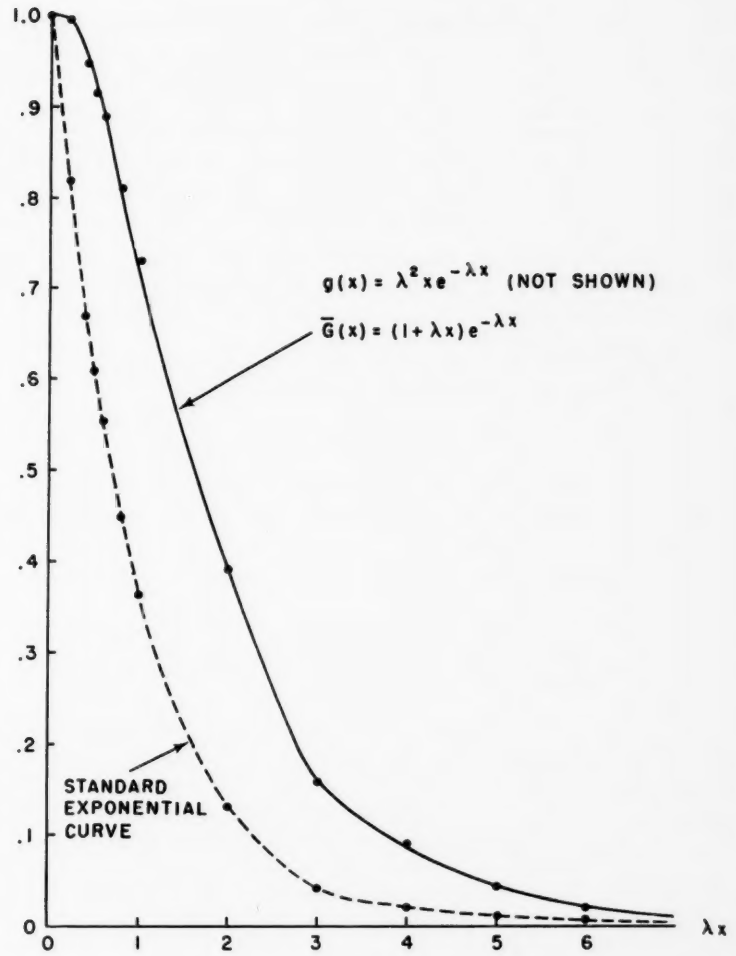

Figure 3. $\bar{G}(x)$ for the distribution of gaps covering an arbitrary time.

Here λ is the arrival rate, as before, but the mean gap size is now $2/\lambda$ instead of $1/\lambda$. $\bar{G}(x)$ is plotted in Figure 3 and contrasted there with the standard exponential curve.

TRAFFIC CHANGING OVER TIME

The discussion up to now has dealt exclusively with the case of completely random arrivals. The only information assumed to be known about these arrivals was λ , their average rate of occurrence over an extended period of time. One interesting characteristic of this simple arrival process is that the probability of an arrival taking place in an infinitesimal interval of time dt is equal to $\lambda\,dt$ no matter where in time dt is located.

An important variation of the completely random process is a partially random process in which arrival times are still independent of each other, but their rate of occurrence changes over time in a known manner. Let the arrival rate be $\lambda(t)$, a given function of time t. The probability now of an arrival taking place in dt is equal to $\lambda(t)dt$ where t is the time at which the infinitesimal interval dt is situated. When $\lambda(t) = \lambda$, a constant over time, the process becomes completely random.

The distribution of gaps generated by the partially random process is said to be non-stationary since it changes with time. The density function of gap x at time t is given by,

$$g(x|t) = \lambda(t + x)e^{-A(t,\,t+x)} \tag{9}$$

where

$$A(t,\ t + x) = \int_{t}^{t+x} \lambda(z)dz. \tag{10}$$

This result holds whether or not t is an ETA; that is, whether x is a gap or a prolongation. The complementary cumulative function is given by,

$$\bar{G}(x|t) = e^{-A(t,\,t+x)} \tag{11}$$

When $\lambda(t) = \lambda$, equations (9) and (11) reduce to equations (1) and (4), respectively, which describe the stationary case discussed originally.

To consider an elementary example of non-stationarity, suppose that at 1:00 p.m. $\lambda(t)$ is 2 planes/hour, it jumps to 3 planes/hour at 2:00 p.m., and it jumps to 4 planes/hour at 3:00 p.m. If t is set at 1:00 p.m., it follows by equation (10) that $A(t,\ t + x)$ increases linearly from 0 to 2 as x increases from 0 hours to 1 hour, and from 2 to 5 as x increases from 1 hour to 2 hours. For x larger than 2 hours,

$$A(t,\ t + x) = 5 + 4(x - 2). \tag{12}$$

By equation (11), this implies that $\bar{G}_1(x|t)$ is as shown in Figure 4. (The subscript is for identification only.)

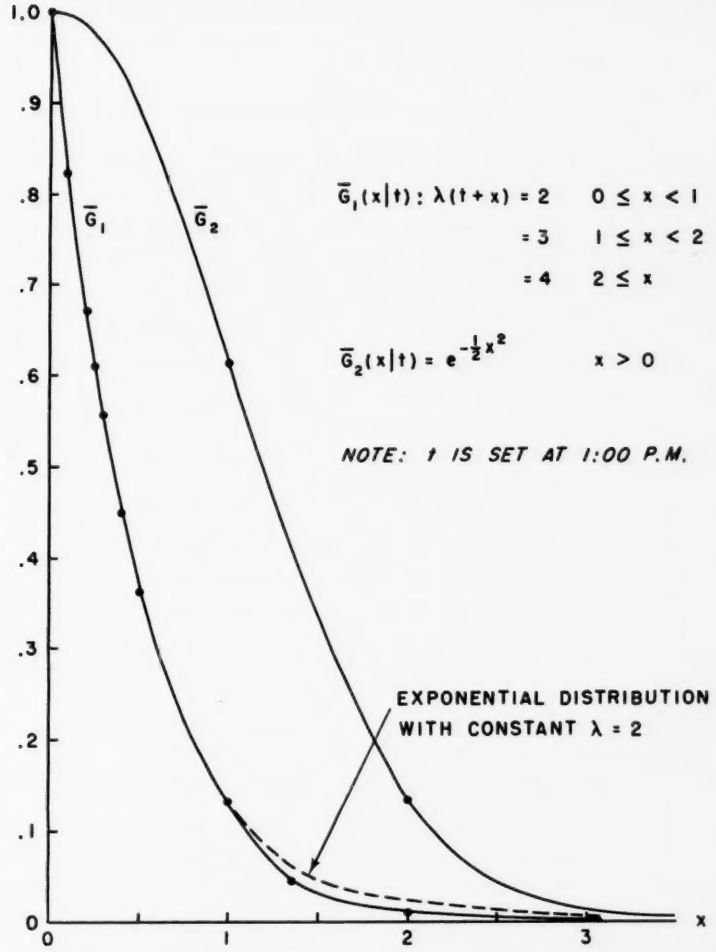

Figure 4. $\bar{G}(x|t)$ for the distribution of gaps in two elementary examples of non-stationarity.

Figure 4 also pictures the exponential curve for a constant $\lambda = 2$. The exponential curve does not seem very different from $\bar{G}_1(x|t)$ for t set at 1:00 p.m., and indeed yields the same samplings for all $u \geq 0.14$. The difference would be more marked, however, if t were permitted to increase beyond 1:00 p.m.

As a second example, suppose that $\lambda(t) = 0$ planes/hour for t set at 1:00 p.m., and thereafter increases linearly as follows:

$$\lambda(t + x) = x \tag{13}$$

where x is the elapsed interval of time since 1:00 p.m. as measured in hours. By equation (10) this implies that,

$$A(t,\ t+x) = x^2/2, \tag{14}$$

and by equation (11) that,

$$\bar{G}_2(x|t) = e^{-x^2/2}. \tag{15}$$

$\bar{G}_2(x|t)$ is also shown in Figure 4. (The subscripts serve to distinguish between the two $\bar{G}$ functions.)

The similarity of form between equations (4) and (11) suggests that an equation analogous to (6) can be used as an alternative to the graphical technique in sampling for x in the non-stationary case, just as (6) provided an alternative in the stationary case. This is indeed so. Letting u be a random drawing from the uniform distribution, as before, we have that,

$$A(t,\ t+x) = \log_e u^{-1} \tag{16}$$

which is an implicit equation for x.

Let us consider the two non-stationary examples in the light of equation (16). In the first example, with the area function $A(t,\ t+x)$ given by equation (12) and the lines preceding, a simple computational procedure for determining x is to let it increase from 0 to that value for which the area function just equals the precomputed value of $\log_e u^{-1}$. In the second example, the procedure is even simpler. With the area function given by equation (14), x may be obtained directly by solving the quadratic equation,

$$x^2/2 = \log_e u^{-1}. \tag{17}$$

DISTRIBUTIONS USED IN THE EN-ROUTE AREA

Having discussed the simulation of plane arrivals and departures at terminals, we turn next to the distributions used in simulating plane arrivals at en-route points, such as reporting fixes and airway intersections. The main difference we encounter is that successive arrivals no longer can be considered independent. In particular, it is not possible for inter-arrival gaps to be near-zero in value. Even though we still work with estimated times of arrival, these ETA's now do reflect scheduling measures; namely, those measures which were exercised by air traffic controllers for the purpose of keeping the planes safely separated at the earlier en-route points.

A certain probability distribution has been found useful in describing the gaps between arrivals which are kept separated in the manner just indicated. If we let s equal the minimum time separation acceptable, and as before call λ the arrival rate (here assumed constant) and x the gap size, then the density function of this distribution may be written,

$$g_1(x) = \frac{\lambda e^{-\lambda(x-s)/(1-\lambda s)}}{1-\lambda s} \qquad x > s \tag{18}$$

The gap size ranges from s to ∞ and the mean gap is $1/\lambda$. The complementary cumulative function is

$$\begin{aligned} \bar{G}_1(x) &= 1 && 0 \le x \le s \\ &= e^{-\lambda(x-s)/(1-\lambda s)} && x > s \end{aligned} \quad (19)$$

(The subscript 1 is again used for identification. This $\bar{G}_1$ should not be confused with the earlier $\bar{G}_1$.)

The distribution represented by function (18) or (19) refers only to gaps. It does not describe prolongations, as did the distributions represented by functions (4) and (11). In the present instance gap sizes cannot be less than s in value, whereas the

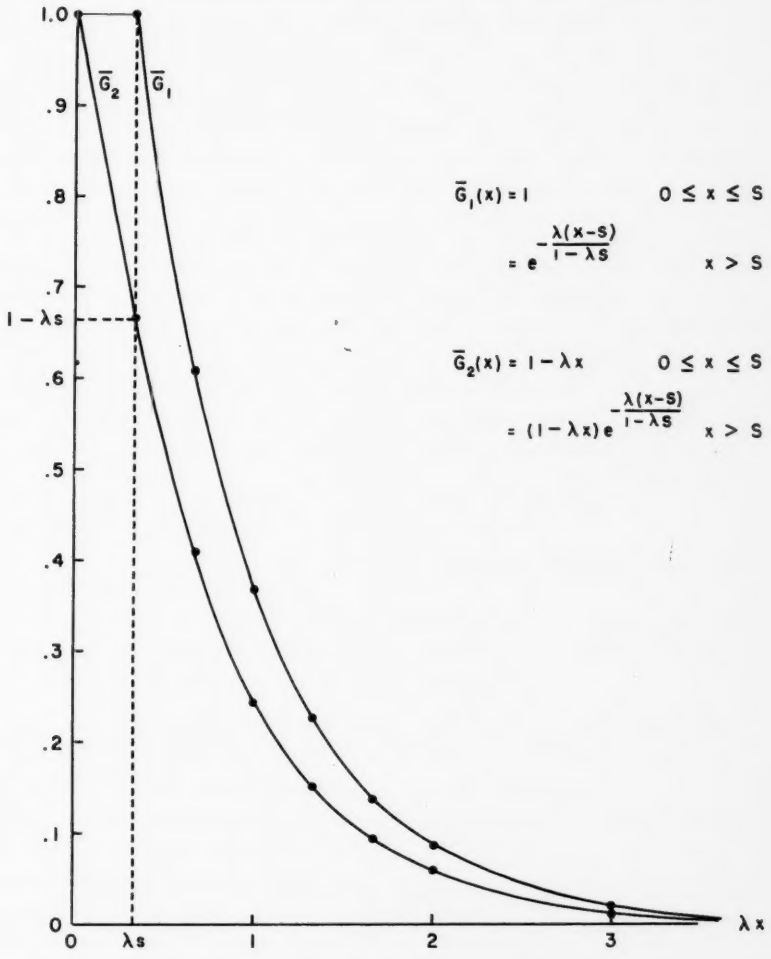

Figure 5. Complementary cumulative functions for the distribution of gaps ($\bar{G}_1$) and prolongations ($\bar{G}_2$) in the en-route area ($\lambda s = \frac{1}{3}$ as an illustration).

lengths of prolongations clearly can approach zero. As a result, the $\bar{G}_1(x)$ of equation (19) is not appropriate for sampling prolongations.

A complementary cumulative function which is appropriate for this purpose is the following:

$$\begin{aligned} \bar{G}_2(x) &= 1 - \lambda x && 0 \le x \le s \\ &= (1 - \lambda s)e^{-\lambda(x-s)/(1-\lambda s)} && x > s \end{aligned} \quad (20)$$

The corresponding density function is,

$$\begin{aligned} g_2(x) &= \lambda && 0 \le x \le s \\ &= \lambda e^{-\lambda(x-s)/(1-\lambda s)} && x > s \end{aligned} \quad (21)$$

The mean gap size is given by,

$$1/\lambda - s(1 - \lambda s/2)$$

which is less than $1/\lambda$ inasmuch as λs cannot exceed 1 by the nature of s. In the en-route case, therefore, the average prolongation is always less than the average gap size, irrespective of the values taken on by λ and s. Moreover, the difference $s(1 - \lambda s/2)$ lies between $s/2$ and s for any λ. The larger λ, the nearer the difference is to $s/2$. The smaller λ, the nearer the difference is to s. For the extreme case of $\lambda = 1/s$, all gaps are exactly s in size, and the prolongations are uniformly distributed between 0 and s. For the other extreme of $\lambda = 0$, all gaps and prolongations are infinite.

Standard regulations for en-route aircraft (not under radar traffic control) call for a minimum acceptable separation of 10 minutes. This corresponds to $s = 1/6$ hour. As an illustration, this s has been combined with $\lambda = 2$ planes/hour to yield $\lambda s = 1/3$, and the complementary cumulative functions shown in Figure 5. $\bar{G}_1(x)$ is used to sample for gaps, and $\bar{G}_2(x)$ for prolongations.

REFERENCES

1. New York Times, December 17, 1960.
2. S. H. Loory, "Electronics, the Human Element, and Air Safety," Reporter Magazine, Feb. 2, 1961
3. E. P. Curtis, Aviation Facilities Planning, Final Report by the Special Assistant to the President, Feb. 10, 1957.
4. Leon Ephross, "Airport Traffic Simulation," Industrial Management Review, MIT School of Industrial Management, April 1960.
5. Leon Ephross, "A Monte Carlo Simulation of Air Traffic Flow into Logan Airport," M.S. Thesis, MIT School of Industrial Management, 1959.
6. Gary Koger, "Transient Analysis of Non-stationary Queuing Systems," M.S. Thesis, MIT School of Industrial Management, 1961.

The Place of a Business Game in the Marketing Curriculum*

by WILLIAM F. MASSY, *Assistant Professor of Industrial Management*, PETER S. KING, *Assistant Professor of Marketing*, and C. DAVIS FOGG, *Graduate Student*

A great deal of interest is currently being devoted to management gaming as a tool for the education of business students and for the development of executives. Some games can be simple, such as Greene and Sisson's Inventory Management Game, which illustrates the square root rule for optimal lot sizes.[1] Other games have attempted to catch the flavor of high pressure sequential decision-making at the general management level. The American Management Association's Top Management Decision Simulation[2] and the IBM Corporation's Management Decision-Making Game[3] are probably the best-known but by no means the only examples. These games require the participating teams to make a series of financial allocations at fifteen minute intervals with information feedback between plays.

More recently, a third type of business game has been developed. Working independently, groups at the Carnegie Institute of Technology[4] and at MIT's School of Industrial Management[5] have

---

[1] Jay R. Greene and Roger L. Sisson, Dynamic Management Decision Games, New York, John Wiley and Sons, 1959.

[2] Franc M. Ricciardi, et al; Top Management Decision Simulation: The A.M.A. Approach; Elizabeth Marting, ed.; American Management Association, New York, 1957.

[3] Discussed in Kalman J. Cohen and Eric Rhenman, "The Role of Management Games in Education and Research," Management Science, Vol. 7, No. 2 (January, 1961), p. 136.

[4] Kalman J. Cohen, et al, "The Carnegie Tech Management Game," Journal of Business, Vol. 33, No. 4 (October, 1960).

[5] P. S. King, W. F. Massy, Arnold E. Amstutz, and Gerald B. Tallman; "The MIT Marketing Game"; Proceedings of the Winter Meeting, December, 1960, American Marketing Association.

* This work was performed in part at the MIT Computation Center, Cambridge, Massachusetts.

developed games that are intended to simulate the broad aspects of a business enterprise and its environment in sufficient detail to allow the players to grapple with a significant and realistic set of business problems. The MIT Marketing Game concentrates upon the problems of product development, distribution, pricing, and sales promotion for a durable good. The CIT game places relatively greater emphasis on the areas of production and finance within the context of the packaged detergent market.

The purpose of this paper is to assess the potential academic usefulness of this third type of business game and then to review the results of the use of the MIT game as the central element in a second year graduate course in marketing management. As the usefulness of the game depends upon the pedagogical objectives it is expected to achieve and the available alternatives for realizing these objectives, it will be necessary to describe briefly (1) the nature of MIT's marketing management course and its relationship to other courses in the marketing curriculum, and (2) other teaching techniques currently employed in teaching marketing management.

ROLE OF THE MARKETING MANAGEMENT COURSE IN THE MARKETING CURRICULUM

MIT's marketing curriculum is divided into three core courses.[6] The introductory course, while adopting the managerial viewpoint, tends to be general, theoretical, and all encompassing. It deals with the environmental elements: the behavior of industrial markets, the structure of trade, and the nature of competition. A textbook, selected readings, and seminar papers provide the basis for effective discussions of contributions from the fields of marketing, economics, sociology, and psychology. Such discussions are relevant for (1) the theory and measurement of consumer demand; (2) the systems approach to the dynamic functioning of distribution channels; and (3) an understanding of the nature of the variables which management manipulates to achieve goals. Case analyses and discussions provide an important supplement to readings, especially in the third area.

The second course, Market Research, emphasizes rigorous problem solving. Study of powerful new analytic techniques is included. But realistic applications of at least some of these techniques are provided through an extensive class project. This combination assures that students achieve a balanced view of the problems of generating information for decision making.

It would be a mistake to assume that students readily learn to apply analytical techniques by working textbook exercises. (They do, of course, learn the mechanics.) The tremendous complexity of marketing decisions requires that the student have some experi-

---

[6] Specialized courses, such as New Product Planning and Development, are also offered.

ence in balancing the risks and advantages of analytical solutions, on the one hand, and of intuition on the other. Knowing how far one can rely on an operations research solution can be one of the more critical questions facing a manager. The student may fail to grasp the relevance of rigorous analytical techniques if they are taught without reference to real situations. He may become frustrated when he realizes that he does not and probably cannot ever know the mathematical form of the advertising effectiveness curve, for example, and may scrap the analytical tools entirely as a result.

Marketing Management, the third basic course, more emphatically than the others, emphasizes the viewpoint of the top marketing decision-maker of the firm. The focus is sharply on realistic problem-solving. Its objective is to familiarize the student with the nature of managerial decision making in the marketing context. To accomplish this objective it is necessary that the student assimilate a considerable amount of information on a wide range of topics. But even more important, he must learn to filter out irrelevant information, to evaluate fragmentary and ambiguous information, and to use this information in the formulation of complex sets of interrelated decisions. The marketing executive must, subject to constraints imposed by the objectives of the firm and by the resource requirements of other units of the firm, accomplish an exceedingly difficult set of tasks. He must develop effective plans based on an analysis of consumer and trade behavior. These plans include price setting, product and product line development, and advertising, sales force, and dealer program development. He must also administer the execution of these plans, modify plans when necessary, and evaluate the consequences of all these actions in order to improve future decisions.

The central objective of the marketing management course at MIT, therefore, is to give students some substantive knowledge of these matters and some understanding of how decision making can be effectively handled in this complex and dynamic environment.

TEACHING TECHNIQUES AVAILABLE

At the present time the great majority of teachers employ the textbook-discussion teaching technique or the case method or a combination of the two in teaching courses in marketing management. Let us assess these basic techniques available for teaching a course of the kind just described.

The textbook-discussion approach has important advantages. It fulfills one important aim of the educational process, the transmission of an accumulated body of knowledge. There are significant hazards, however, in the exclusive use of this technique. In the first place, it is difficult to classify business phenomena precisely, and often next to impossible to communicate the real meaning of these general classifications to the uninitiated. There is danger that the student will associate otherwise valid generalizations with the wrong situations; he may miss the essential and often subtle

differences that set one problem apart from another. Second, if the student comes "to revere the principles as eternal verities, this can easily lead to a closed system of reasoning that impedes or prevents the adventurous exploring of new possibilities and the abandoning of outmoded notions or customs."[7] Third, it puts a heavy premium on the assimilation of factual information. Clearly there is more to effective and imaginative marketing management than a long memory and a fast slide rule. The crucial limitation of the method, however, is that the technique is simply not a sufficient vehicle for imparting the essence of the management experience.

The case method of instruction has been one of the landmarks in the history of business education. Case study attempts to teach practical problem solving by allowing the student to project himself vicariously into a business situation and react, as in the real world, to the multitude of pressures bearing upon his decisions. While we all recognize its power, case analysis has its limitations, too. It is often difficult to motivate members of the class to sift and to apply their fund of knowledge systematically. Failure to sift one's knowledge and failure to set adequate goals can lead to the generation of "quick" solutions based more on intuition than is necessary, given the facts. It is possible to argue that the necessity of presenting and defending a position in a vigorous class discussion insures that the case analyst will be strongly ego-involved. We submit that although vigorous discussions take place, they may be superficial.

Contrary to common opinion, the case method is a rather rigid and static teaching technique. Students cannot clarify the information presented or effectively explore courses of action for which no data are available. They must focus on a specific problem or set of problems because of lack of information. As a result students tend to suboptimize simply because they have never had a clear view of the firm as a whole. Finally, and perhaps most importantly, the sequential character of decision making is neglected in case analysis. The student gets only a single shot at the problem. He is always placed in the position of taking a quick look at a situation with which he has no involvement and of making recommendations without being fully aware of all the pressures that bore upon the managers who guided their firm into the case situation. He never has the opportunity to study and evaluate the consequences of his recommendations or the opportunity to correct his own mistakes.[8]

---

[7] Frank G. Pierson, et al, The Education of American Businessmen, McGraw-Hill Book Co., New York, 1959, page 440.

[8] The Dynamic Business Case, recently developed by the Swedish Research Institute, appears to be free of some of these defects. The new method has much in common with the CIT-MIT type of business game. It is described in: Cohen and Rhenman, op. cit.

The business game as a tool for teaching marketing management. The authors believe that many of the difficulties cited above can be overcome through the use of an appropriate business game, supplemented by case discussions and readings. A game provides for a sequence of decisions, introduces dynamic variables, allows a greater degree of familiarity with the environment through continued information feedback and market research, and if handled properly can generate a substantial degree of ego-involvement. On the negative side, however, lie the rather impressive capital and operating costs, not to mention the large amounts of time students typically put into the preparation of decisions. Also, there is the question of what the feedback itself teaches the students: do they tend to learn falsehoods or to generalize inappropriately from the reaction functions built into the game? Will they come to anticipate that the real world is as structured a place as the game environment? Or, if the game is a highly detailed simulation such as the MIT game, will the students become overwhelmed by complexity and will their decisions degenerate to the level of mere guesses?

The answers to these queries are not yet clear, though it is obvious that they depend strongly upon the particular game utilized. We believe the difficulties can be overcome. A number of tentative conclusions which have been suggested by our experience with the MIT Marketing Game are presented in the next section and are illustrated by a description and evaluation of one series of plays.

To be an effective teaching tool within the context of an advanced marketing management course, as envisioned above, a game must be detailed enough to allow decisions to be made on a wide range of marketing variables. Our goal is not to "stimulate an attitude of experimentation with respect to the environment," nor to "generate an appreciation of the interrelations among decision areas," nor even "to have a good time," as has sometimes been the case for the general management games of the AMA type.[9] While each of these is an important by-product of any business game play, we emphasize that the sine qua non for the use of a game in the advanced marketing management course is that the exercise contributes more effectively than other teaching techniques to the understanding and use of substantive marketing tools. It is not enough, for example, that the student acquire "some feel" for balancing financial allocations for sales calls and advertising within the limits of the cash budget of the firm. If he is to gain an understanding of the interrelationship between these tools, he must at the very least grapple with the problems of selecting promotional appeals, of training and motivating the sales force, and of stimulating the trade. He must learn to consider his advertising budget in terms of the number of persons reached through media such as Life magazine or network television, rather than merely as a number of homogeneous dollars. Even the type of show or copy content

[9] Ibid., p. 144.

is highly relevant in evaluating the impact of many advertising appeals. The case method currently affords a way of dealing with these problems. If the inherent dynamic advantages of business gaming are to make themselves felt, the game must allow their consideration also.

The MIT Marketing Game is, to the best of our knowledge, the only game presently available which encompasses this necessary range and detail of marketing variables. Many operational problems remain to be solved, but the results gained so far are highly encouraging.

During the construction and testing of the game, we evolved the following "principles" upon which an effective marketing game must be based:

(1) In view of the richness of decision alternatives necessary for the effective teaching of marketing management, the game should follow the model of "Free Kriegspiel" rather than of "Rigid Kriegspiel," the two types of war gaming that grew up in Prussia during the nineteenth century.[10] In the free variety of war game, the experienced judgment of human referees played an important role. In developing a comprehensive marketing game, one must similarly make provision for referees to rate variables such as advertising copy and appeals. Use of referees avoids the need to design an impossibly elaborate model and simultaneously the need to restrict the freedom of action of the players. Referees are thus an essential feature of a comprehensive marketing game.

(2) Reality, freedom of student action, sensitivity to decision, and general plausibility of results are the most important characteristics of the marketing game. A particular real industry must be simulated with sufficient accuracy so that empirical data on product characteristics, known consumer preferences, media advertising rates and coverage, available types of outlets and their characteristics, and economic considerations can be open for student research. Once convinced that arbitrariness and artificiality have been removed, students can be encouraged to place themselves in the role of marketing managers of real firms. Achieving this frame of mind is, of course, one of the central objectives of the comprehensive game. Students come to feel responsible for the fortunes of their firms. They are competing with their peers, not, as in a case analysis, attempting to improve on the decisions of experienced businessmen. They are made to feel that they have realistic freedom to respond to market vicissitudes and that operating results are meaningful and fair. The instructor becomes a con-

---

[10] Cf. Ibid., p. 132.

sultant to men with pressing and difficult problems. In these circumstances student identification with the role of the marketing manager comes readily and easily.

(3) In formulating detailed, realistic programs the student must have access to information equivalent to that which would be available in the real world. Most games include provision for market research and experimentation within the context of the game structure. When referees are used they, too, provide exceedingly valuable feedback data. They are particularly needed in providing feedback of qualitative information and are the source of all quantitative feedback not generated by the computer. We believe, in addition, that the student should be encouraged to reason by considering marketing principles and analogies from the real world. If the student views the game as realistic, supplementary readings become aids in solving immediate problems and not merely problems he may encounter ten years hence. Thus, the student, because he is playing the game, may become more, rather than less, interested in supplementary readings and cases. The use of real world information need not diminish the significant aspects of the structure of the environment simulation. But how different this is from the blind probing which would result from students being forced to assume the real world offers few lessons for the game player.

(4) Even though one is attempting to construct a realistic game in the sense described above, it will be necessary to abstract and simplify, especially within the mathematical model. The abstractions should be performed so that it is worthwhile for the student to apply systematic techniques and valid marketing principles. It may be practical, for example, to construct an advertising effectiveness function in such a way that a team, which intelligently uses linear programming to find a media plan which maximizes exposure among a certain population, will ceteris paribus do better than a competitor who does not. The MIT Marketing Game cannot at present boast such a result.

In summary, the authors view the comprehensive marketing game as a vehicle through which the use of marketing tool variables can be taught. The game itself cannot do the job, as the simulation is not completely accurate and the outputs are often as difficult to interpret as events in the real world itself. In the last analysis, the actual learning and teaching responsibilities lie where they always have - with the student as he studies the literature and interacts with his fellows and with the instructor as he guides the students' efforts and interprets his own experience. The game merely provides a focal point for their efforts.

Description of the MIT Marketing Game. The MIT Marketing Game utilizes student teams who operate competing companies in a mock national floor waxer market. The market itself is simulated by a mathematical model programmed for the IBM 650 and 709 computers. (The important features of the model are presented in the appendix.) The game is structured so that decisions may be made for each 3, 6, or 12 month period. In the most recent use of the game, each play represented 6 months of the company's history. Thus in executing six plays the students ran their companies for a total of three years. Each play of the game took approximately one week of real time.

Prior to the start of the game play, the teams are given data representing several periods of their company's history. Therefore, the students have an opportunity to study the company's financial history, marketing mix, and distribution and sales policies before making their first series of decisions. It generally takes this initial review plus one or two actual plays to familiarize the teams with the market and with the mechanics of game play.

Although the primary emphasis of the game is the determination of the marketing mix for each decision making period, the game forces the student to consider virtually all of the important variables surrounding decision making in a going concern. Each team, within the limitations of its company's financial structure, must define its long range goals. They then make the period-by-period decisions concerning their short run goals and the tactics to be utilized in reaching them. As in the real world, both short and long run goals and tactics are constantly reviewed and modified as the market situation changes.

Among the quantitative decisions required each period, teams must determine how much money is available to them and then disburse what is needed into advertising, sales force expense, maintenance of plant capacity, building of new plant capacity, debt retirement, dealer promotion expenses, market research, product research, process research, inventory maintenance, etc. Prices must be set and distribution channels must be determined. The teams have the option of selling direct or utilizing wholesalers and must determine appropriate margins and advertising allowances for each level of distribution. The teams may sell in a multitude of types and sizes of outlets as well as in any or all of the country's three economic regions (North, South, and West). The companies soon find themselves deeply involved in the question of sales forecasts, expected distribution of sales, budget allocations, and alternative tactics and goals. The teams, in short, must design, within their individual financial restraints, a marketing mix that they feel will be most effective in reaching the company's goals. The budget allocations and cash flow figures that come out of this section of the analysis constitute the teams' quantitative inputs.

In order to ease the administrative burden and keep attention focused on the marketing aspects of the game, certain assumptions and restrictions are given. First, a company may market either a standard model waxer or a deluxe model (heavier motor and construction, more features) in any of the three economic regions of the country. Thus one may sell the standard model in one region and the deluxe model in the other two, but not both in the same region. This limitation, of course, makes the marketing of a product line impossible and does, we feel, take away one of the primary marketing tools found in the real world. It is one of the major faults of the game as presently constituted.

Second, certain fixed and variable costs are given for each company. Administrative burden and unit labor, material, inventory holding, and transportation costs are fixed. The cost of maintaining a company salesman is given, as are the rates for various advertising media, certain point of sale materials, market research surveys of various kinds, etc. All of these figures are based on the probable costs as found in the real world. If a strategy involves costs which are not given in advance, the team prepares a proposal stating and justifying the figures involved. The proposal is then approved or modified by the referees.

In addition to the quantitative factors, the MIT game allows consideration of the qualitative facets of marketing. Each competing team has complete freedom to present detailed programs elucidating the various policies and strategies to be used in the play. For example, the teams present their advertising appeals and strategy by writing the copy and structuring the format of the ads which are placed in the media of their choice - regionally and/or nationally, depending upon actual availability. In a similar manner, plans are designed for dealer promotion, point of sale activity, product design, sales force compensation and motivation, sales call policy, special promotions, etc. Thus, the student refines and implements the marketing methods to be used. He is thereby motivated to delve deeply into the theory and practice of the marketing tools at his command.

These detailed plans, which we may call qualitative inputs, along with the quantitative inputs discussed above become the team's program for the period in question. The qualitative inputs are submitted to a control team of referees made up of several professors, who evaluate the detailed strategies of the competing teams and assign "effectiveness" ratings to their various budgetary allocations. For example, Company A may budget $100,000 for media advertising directed to consumers. The control team, reacting as consumers, may consider the appeals, layout, copy, and media placement relatively ineffective. Therefore, the control team may decide that Company A's $100,000 will, actually, be worth only $70,000. Thus for every dollar spent, Company A's brand receives only $0.70 worth of consumer advertising. In addition, it may be decided that Company A's appeals will stimulate

a great deal of primary demand (demand for floor waxers in general) as opposed to selective demand (demand for Company A's brand). The control team may therefore decide that 50 percent ($35,000) of Company A's advertising dollars will go toward stimulating industry demand while the remaining $35,000 will be effective in bringing consumers into stores to see Brand A. In this manner, the control team assigns effectiveness ratings to each applicable quantitative input figure.

The assembled data for the competing teams is fed into the computer and the play completed. It should be noted that the economic conditions surrounding the market at the time of the play can be controlled by the game administrator. This added bit of flexibility means that the students may find themselves coping with a recession at any time.

At the end of each play, each company receives its income statement, balance sheet, company and industry sales data broken down by region, and a breakdown of company sales by store type, size, and region. Each company is supplied with such information about its competitors as would normally be available in the real world, i.e., the type of outlets in which they are selling, the content of their advertising and point of sale material, their retail prices, their product design, and a rough estimate of their advertising and point of sale budgets. At the end of every year of play, each team receives its competitor's annual income statement and balance sheet. Additional information about competitors, consumers, and the market in general can be explored, for a price, by conducting market research.

Market research information is provided by a game administrator. The students must structure the research methods and the questions to be directed at the consumer or the trade. The better the structure of the research, the more reliable the answers, so the students are motivated to gain an understanding of the techniques and methodology of market surveys and to be on guard against the bias inherent in such procedures. In this manner, the teams may gain a more complete perception of competitive tactics, the effectiveness of various promotional techniques, and the nature of the consumer and the market place.

Some of the feedback information is realistically incomplete. The student knows little more about the financial affairs of his competitor than would the informed stockholder. Knowledge of the consumer and of the trade is improved with research and experience. The student is forced to make decisions with incomplete information. He cannot avoid the necessity for making a decision (as can be done in the case method) by saying, "If I had more information about this I...." He has the choice of either getting the information - time, money, and methodology permitting - or doing without it.

THE MIT MARKETING GAME IN ACTION

Further light can be thrown upon the nature of the educational process inherent in game play by assessing some aspects of the history of an actual play. In this instance, the game was being used in a second year graduate course in marketing management conducted during the Fall of 1960. The pages that follow contain a description of an important strategic decision and comments about the use and value of certain portions of the exercise.

A Team Strategy. At the outset of the game, Company A found itself holding a 35% share of the floor waxer market, its sole competitor having captured the remainder. The Company's financial position was poor. It was saddled with a high amount of debt, low borrowing power, and little cash. On the other hand, the Company's competitors found themselves in excellent financial shape. Nevertheless Company A decided that its primary goal was "to seize a commanding share of the market." It was believed that the waxer market was largely undeveloped and that media advertising would contribute more heavily to primary, rather than selective, demand. By keeping its own media advertising expenditures low and by emphasizing point of sale promotion, Company A hoped to divert consumers from the competitor's product.

The general strategy proved successful. Sales increased rapidly in the successive three periods and Company A achieved a commanding share of the market. A lag in the liquidation of accounts receivable, however, produced insufficient additional cash flow to meet the increased production costs involved with the expanded sales. An additional difficulty arose when Company B cut prices severely in one of the three regions.

A reevaluation of long-run goals was required of Company A. Possession of the major share of the market had increased rather than decreased the financial pressures on the firm. In addition, sufficient development had occurred in the market to require Company A to assume its share of media advertising if it intended to maintain its sales. Furthermore, it was clear that the competitor was in an excellent position to indulge in severe price competition. The company could not maintain sales at the current rate without endangering profits in the near future.

Under these pressures Company A decided that a majority share of the market was not a desirable short run criterion. Instead, regaining industry sales leadership was established as the long run goal, with profitability the dominant short run criterion of performance. Therefore, Company A concluded that, if necessary, it would sacrifice part of its market share in order to protect profits and to improve the firm's financial condition. In addition, Company A felt compelled to insulate itself from severe price competition with Company B.

Product differentiation appeared to offer a solution to the problem; a deluxe model polisher featuring a variable speed motor was designed. It was priced $20 above the standard model being retailed by the competitor in two regions, and $10 above the deluxe model marketed in a third.

Because it was believed that dealer push was essential to the sale of such a high-priced product, margins and sales calls to dealers were increased, additional push money (spiffs) was given to retail salesmen, and a selective distribution policy was solidified to concentrate sales in compatible outlets of high potential. Realizing that competitors would, in the next period, retaliate either by cutting price further (which was considered unlikely) or by emulating A's product features, Company A initiated an extensive consumer advertising campaign to create a favorable brand image for its product. The image desired for Company A's waxer was that of high quality, versatility, and ease of use. In addition, a long term warranty was introduced to back up the expected favorable price-quality association. The variable speed feature was dramatized by a descriptive and catchy name, Selectramatic, and high impact advertising was designed to convince the consumer of the need for the variable speed feature.

Company A's share of the market, in units, dropped from nearly 50% to around 35%, but its profits more than doubled. (Company A maintained industry profit leadership throughout the remainder of the game.) When an imported Japanese deluxe model, priced $20 below Company A's brand, was introduced on the fifth play, Company A lost only $2\frac{1}{2}\%$ of the market while the Japanese firm took nearly 33% away from the other American competitor.

CONCLUSION

The example shows the realism involved in the game and the detail and thought that is elicited from the student. The profit and loss statement is real and relevant, and this fact is never forgotten. However, the financial facts of the market place do not necessarily breed conservatism. On the whole, tactics were carefully thought out, and team goals and strategies were usually kept in mind.

An important question arises as to the extent of student recognition of the need for sequentially consistent decisions. We would say that the problem was recognized in varying degrees by the four teams participating in this Fall's game. For example, Company A, when making the decision to market the deluxe model, formulated plans, including alternatives for several eventualities, for several periods in advance. In other instances the Company realized the value of stability from the retailer's point of view and resisted any temptations to juggle prices and margins from period to period.

On the other hand, some teams would reverse policies from period to period as expediency and the current state of competition seemed to dictate. This juggling could, if taken to its logical ex-

treme, result in each period being looked upon as a separate case. This is an undesirable viewpoint, and measures aimed at avoiding it need to be implemented in future class discussions.

In planning his marketing strategy, the student finds himself speculating on the various structural elasticities and lags inherent in the market. It was often very difficult to isolate the cause or causes of any particular success or failure, much less separate out the effect of particular efforts by the firm or its competitors on sales volume and distribution. The students, however, proved to be quite astute in making these analyses even with extremely incomplete information. By combining available feedback data and market research information with a good theoretical background in marketing, teams felt that they were able to determine with rough accuracy the various elasticities and elements important to their particular program.

The analytical techniques used were not particularly profound and, indeed, there was a strong belief that the time available and the nature of the feedback information did not make themselves amenable to sophisticated techniques. The students did make considerable use of rough graphical and tabular methods of summarizing and interpreting data. At least one team kept a running graphical record of company and industry sales by region and then attempted to explain changes by noting any changes in the content of their marketing mix. More work on this point is clearly necessary.

Certain written assignments based upon the game were utilized. The students were required to describe and analyze their marketing strategies and tactics so that the instructor could judge the quality of their work. At the end of the course each team member wrote a semi-annual report for his company outlining the important facts and events for a particular period. The reports were duplicated and collected to form a detailed summary of the company's performance. Finally, all the students were given copies of their own and their competitors' reports and were asked to write an analytical history of the industry. The assignment was not only valuable in crystallizing the experience gained by the student, but it provided valuable information for the game administrators as well.

Time commitments were heavy for all involved in the game. The enthusiasm of the students meant that the professors teaching the course sections had to be readily available for consultation, and the administration of the game itself occupied the full time of a graduate research assistant. For the student, it meant six to ten hours per week in preparation and team discussion in addition to the normal reading, case assignments, and three regular hours of class time. But regardless of the burdens involved, the twenty-four graduate students who participated in the game during the Fall of 1960 almost unanimously agreed that it was one of the more interesting and valuable experiences of their business education.

Our game will never supplant an intensive basic course in marketing. The theoretical premises supplied by case-discussion-text methods are the foundation upon which the student builds as he experiences the game simulation. Here at MIT the game playing itself was interspersed with lectures, cases, and textual material that dealt with the situations encountered in the game. At least one class hour each week was reserved for cases drawn from other industries and new situations. It should be noted that the teams expended a great deal of time and effort in voluntarily searching the literature for data and ideas.

The work reported in this paper is naturally just a beginning. We are learning more about the capabilities and limitations of our game in each class for which it is used. Administrative techniques have had to be devised and constantly revised in order to handle the flow of information to and from the student teams. A systematic effort to define and evaluate what the student gains through taking the game-oriented course is under way, and a revision of the existing game's mathematical model to incorporate some of the functions previously handled by the human referees is scheduled for the near future.

Of great promise, also, is the recent interest in gaming on the part of faculty members from the areas of business economics, production, accounting, and finance within the School of Industrial Management. If agreement can be reached about the teaching goals to which gaming can contribute in these areas, MIT can look forward to the development of a Comprehensive Management Game in the near future.

APPENDIX

DESCRIPTION OF THE MODEL USED IN THE MIT MARKETING GAME

The mathematical model is split into two sections. The market simulation section computes industry and company sales for each of the competing companies in each region. The second section takes the sales data computed by the simulation model and, in combination with the assigned fixed and variable costs and budgeted expenditures, computes a profit and loss statement and balance sheet for each company. The programming of the second section is in accordance with standard accounting procedures and therefore needs no discussion.

Working in response to economic factors, initial market conditions, and total marketing activity during the period, the market simulation model first generates total industry sales by region. It then computes each company's sales in each region by considering such competitive factors as the price, quantity, and quality of advertising, sales force activity, product quality, etc., for each competitor.

The generation of industry sales is governed by reference to a pattern of normal industry sales expressed as a percent of families that will purchase a waxer at any point in the sales history of the market. This pattern indicates that acceptance of a new appliance, while slow at first, rapidly increases once a certain level of market saturation is reached. The curve, of course, tapers off as full saturation is approached. The normal sales pattern provides for seasonal and cyclical variations, and the three regions are at different levels of saturation at the beginning of the game.

Actual industry sales are determined from the normal sales as influenced by elasticity functions accounting for all competitors' product prices and qualities, sales promotion activities, adequacy of distribution, and retail margins. These functions have the expected "S" shape: that is, there are levels above and below a normal range where one would expect marginal dollars to be less effective.

Each company's sales are in turn a function of the competitiveness of its prices, distribution policies, and advertising. Because the operation of this part of the model is quite complex, there is considerable interaction between competitors as well as among the individual elements of a single company's marketing mix. One example is the distribution function, (i.e., the competitive effectiveness of distribution policies), which depends on the interaction of the number of sales calls on a store, the amount and quality of point of sale materials and trade advertising, and the level of retail margins and advertising allowances. A "dealer enthusiasm" function takes into account the past selling success, in previous periods, of the brand in a particular outlet as well as the loss of retailer interest to be expected if the company has selected a set of incompatible types of outlets (e.g., auto accessory stores competing with large department stores). In addition, the effectiveness of sales calls is related to the level of media advertising expenditures.

Each of the aforementioned functions have threshold and saturation values. For example, it will take a certain level of selling activity to induce a dealer just to stock the product, much less to push it. On the other hand, marginal dollars spent beyond a certain sales activity level add little to the selling effort expended by the dealer.

In addition to these elasticities, the simulation model has certain built-in lags that determine the time at which previously introduced elements of the marketing mix become effective. For example, with a given level of promotional activity, when a new outlet is added it may take several months for the volume of sales to reach the expected level. On the other hand, the effective gains of a magazine advertising campaign may continue for some time after the campaign is dropped. The model has an implicit decay constant to reflect this effect. In a like manner, lags and decay constants are built into the model to reflect the timing and ef-

fectiveness of changes in the level of newspaper advertising expenditure, sales call frequencies, number of outlets, and so on.

As the previous discussion shows, the authors of the game took great pains to assure a realistic simulation of the market. While the elasticity functions may not coincide exactly with those found in the real world, they are realistic enough to confront the student with a market that is as sensitive and responsive as that of the real world.

Management Development as a Process of Influence*

by EDGAR H. SCHEIN, *Associate Professor of Industrial Management*

The continuing rash of articles on the subject of developing better managers suggests, on the one hand, a continuing concern that existing methods are not providing the talent which is needed at the higher levels of industry and, on the other hand, that we continue to lack clear-cut formulations about the process by which such development occurs. We need more and better managers and we need more and better theories of how to get them.

In the present paper I would like to cast management development as the problem of how an organization can influence the beliefs, attitudes, and values (hereafter simply called attitudes) of an individual for the purpose of "developing" him, i.e. changing him in a direction which the organization regards to be in his own and the organization's best interests. Most of the existing conceptions of the development of human resources are built upon assumptions of how people learn and grow, and some of the more strikingly contrasting theories of management development derive from disagreements about such assumptions.** I will attempt to build on a different base: instead of starting with assumptions about learning and growth, I will start with some assumptions from the social psychology of influence and attitude change.

Building on this base can be justified quite readily if we consider that adequate managerial performance at the higher levels is at least as much a matter of attitudes as it is a matter of knowledge and specific skills, and that the acquisition of such knowledge and skills is itself in part a function of attitudes. Yet we have given far more attention to the psychology which underlies change in the area of knowledge and abilities than we have to the psychology

---

* I am greatly indebted to Warren Bennis and Douglas McGregor, whose helpful comments on the first draft of this paper have helped me to refine many of the ideas in it.

** An excellent discussion of two contrasting approaches - the engineering vs. the agricultural - deriving from contrasting assumptions about human behavior can be found in McGregor, 1960, Chapter 14.

which underlies change in attitudes. We have surprisingly few studies of how a person develops loyalty to a company, commitment to a job, or a professional attitude toward the managerial role; how he comes to have the motives and attitudes which make possible the rendering of decisions concerning large quantities of money, materials, and human resources; how he develops attitudes toward himself, his co-workers, his employees, his customers, and society in general which give us confidence that he has a sense of responsibility and a set of ethics consistent with his responsible position, or at least which permit us to understand his behavior.

It is clear that management is becoming increasingly professionalized, as evidenced by increasing emphasis on undergraduate and graduate education in the field of management. But professionalization is not only a matter of teaching candidates increasing amounts about a set of relevant subjects and disciplines; it is equally a problem of preparing the candidate for a role which requires a certain set of attitudes. Studies of the medical profession (Merton, Reader, and Kendall, 1957), for example, have turned their attention increasingly to the unravelling of the difficult problem of how the medical student acquires those attitudes and values which enable him to make responsible decisions involving the lives of other people. Similar studies in other professions are sorely needed. When these are undertaken, it is likely to be discovered that much of the training of such attitudes is carried out implicitly and without a clearly formulated rationale. Law schools and medical schools provide various kinds of experiences which insure that the graduate is prepared to fulfill his professional role. Similarly, existing approaches to the development of managers probably provide ample opportunities for the manager to learn the attitudes he will need to fulfill high level jobs. But in this field, particularly, one gets the impression that such opportunities are more the result of intuition or chance than of clearly formulated policies. This is partly because the essential or pivotal aspects of the managerial role have not as yet been clearly delineated, leaving ambiguous both the area of knowledge to be mastered and the attitude to be acquired.

Existing practice in the field of management development involves activities such as: indoctrination and training programs conducted at various points in the manager's career; systematic job rotation involving changes both in the nature of the functions performed (e.g. moving from production into sales), in physical location, and in the individual's superiors; performance appraisal programs including various amounts of testing, general personality assessment, and counseling both within the organization and through the use of outside consultants; apprenticeships, systematic coaching, junior management boards, and special projects to facilitate practice by the young manager in functions he will have to perform later in his career; sponsorship and other comparable activities in which a select group of young managers is groomed system-

atically for high level jobs (i.e. made into "crown princes"); participation in special conferences and training programs, including professional association meetings, human relations workshops, advanced management programs conducted in business schools or by professional associations like the American Management Association, regular academic courses like the Sloan programs offered at Stanford and MIT, or liberal arts courses like those offered at the University of Pennsylvania, Dartmouth, Northwestern, etc. These and many other specific educational devices, along with elaborate schemes of selection, appraisal, and placement, form the basic paraphernalia of management development.

Most of the methods mentioned above stem from the basic conception that it is the responsibility of the business enterprise, as an institution, to define what kind of behavior and attitude change is to take place and to construct mechanisms by which such change is to occur. Decisions about the kind of activity which might be appropriate for a given manager are usually made by others above him or by specialists hired to make such decisions. Where he is to be rotated, how long he is to remain on a given assignment, or what kind of new training he should undertake, is masterminded by others whose concern is "career development." In a sense, the individual stands alone against the institution where his own career is concerned, because the basic assumption is that the institution knows better than the individual what kind of man it needs or wants in its higher levels of management. The kind of influence model which is relevant, then, is one which considers the whole range of resources available to an organization.

In the remainder of this paper I will attempt to spell out these general themes by first presenting a conceptual model for analyzing influence, then providing some illustrations from a variety of organizational influence situations, and then testing its applicability to the management development situation.

A MODEL OF INFLUENCE AND CHANGE

Most theories of influence or change accept the premise that change does not occur unless the individual is motivated and ready to change. This statement implies that the individual must perceive some need for change in himself, must be able to change, and must perceive the influencing agent as one who can facilitate such change in a direction acceptable to the individual. A model of the influence process, then, must account for the development of the motivation to change as well as the actual mechanisms by which the change occurs.

It is usually assumed that pointing out to a person some of his areas of deficiency, or some failure on his part in these areas, is sufficient to induce in him a readiness to change and to accept the influencing agent's guidance or recommendations. This assumption may be tenable if one is dealing with deficiencies in intellectual skills or technical knowledge. The young manager can see, with

some help from his superiors, that he needs a greater knowledge of economics, or marketing, or production methods, and can accept the suggestion that spending a year in another department or six weeks at an advanced management course will give him the missing knowledge and/or skills.

However, when we are dealing with attitudes, the suggestion of deficiency or the need for change is much more likely to be perceived as a basic threat to the individual's sense of identity and to his status position vis-a-vis others in the organization. Attitudes are generally organized and integrated around the person's image of himself, and they result in stabilized, characteristic ways of dealing with others. The suggestion of the need for change not only implies some criticism of the person's image of himself, but also threatens the stability of his working relationships because change at this level implies that the expectations which others have about him will be upset, thus requiring the development of new relationships. It is not at all uncommon for training programs in human relations to arouse resistance or to produce, at best, temporary change because the expectations of co-workers operate to keep the individual in his "normal" mold. Management development programs which ignore these psychological resistances to change are likely to be self-defeating, no matter how much attention is given to the actual presentation of the new desired attitudes.

Given these general assumptions about the integration of attitudes in the person, it is appropriate to consider influence as a process which occurs over time and which includes three phases:

(1) Unfreezing:* an alteration of the forces acting on the individual, such that his stable equilibrium is disturbed sufficiently to motivate him and to make him ready to change; this can be accomplished either by increasing the pressure to change or by reducing some of the threats or resistances to change.

(2) Changing: the presentation of a direction of change and the actual process of learning new attitudes. This process occurs basically by one of two mechanisms: (a) identification** - the person learns new attitudes by identifying with and emulating some other person who holds those attitudes; or (b) internalization - the person learns new attitudes by being placed in a situation where new attitudes are demanded of him as a way of solving problems which confront him and which he cannot avoid; he discovers the new attitudes essentially for himself, though the situation may guide him or make it probable that he will discover only those attitudes which the influencing agent wishes him to discover.

---

* These phases of influence are a derivation of the change model developed by Lewin (1947).

** These mechanisms of attitude change are taken from Kelman (1958).

(3) Refreezing: the integration of the changed attitudes into the rest of the personality and/or into ongoing significant emotional relationships.

In proposing this kind of model of influence we are leaving out two important cases - the individual who changes because he is forced to change by the agent's direct manipulation of rewards and punishments (what Kelman calls "compliance") and the individual whose strong motivation to rise in the organizational hierarchy makes him eager to accept the attitudes and acquire the skills which he perceives to be necessary for advancement. I will ignore both of these cases for the same reason - they usually do not involve genuine, stable change, but merely involve the adoption of overt behaviors which imply to others that attitudes have changed, even if they have not. In the case of compliance, the individual drops the overt behavior as soon as surveillance by the influence agent is removed. Among the upwardly mobile individuals, there are those who are willing to be unfrozen and to undergo genuine attitude change (whose case fits the model to be presented below) and those whose overt behavior change is dictated by their changing perception of what the environment will reward, but whose underlying attitudes are never really changed or refrozen.

I do not wish to imply that a general reward-punishment model is incorrect or inappropriate for the analysis of attitude change. My purpose, rather, is to provide a more refined model in terms of which it becomes possible to specify the differential effects of various kinds of rewards and punishments, some of which have far more significance and impact than others. For example, as I will try to show, the rewarding effect of approval from an admired person is very different in its ultimate consequences from the rewarding effect of developing a personal solution to a difficult situation.

The processes of unfreezing, changing, and refreezing can be identified in a variety of different institutions in which they are manifested in varying degrees of intensity. The content of what may be taught in the influence process may vary widely from the values of Communism to the religious doctrines of a nun, and the process of influence may vary drastically in its intensity. Nevertheless there is value in taking as our frame of reference a model like that proposed and testing its utility in a variety of different organizational contexts, ranging from Communist "thought reform" centers to business enterprises' management development programs. Because the value system of the business enterprise and its role conception of the manager are not as clear-cut as the values and role prescriptions in various other institutions, one may expect the processes of unfreezing, changing, and refreezing to occur with less intensity and to be less consciously rationalized in the business enterprise. But they are structurally the same as in other organizations. One of the main purposes of this paper, then, will be to try to make salient some features of the influence of the

organization on the attitudes of the individual manager by attempting to compare institutions in which the influence process is more drastic and explicit with the more implicit and less drastic methods of the business enterprise.

ILLUSTRATIONS OF ORGANIZATIONAL INFLUENCE

Unfreezing. The concept of unfreezing and the variety of methods by which influence targets can be unfrozen can best be illustrated by considering examples drawn from a broad range of situations. The Chinese Communists in their attempt to inculcate Communist attitudes into their youth or into their prisoners serve as a good prototype of one extreme. First and most important was the removal of the target person from those situations and social relationships which tended to confirm and reinforce the validity of the old attitudes. Thus the targets, be they political prisoners, prisoners of war, university professors, or young students, were isolated from their friends, families, and accustomed work groups and cut off from all media of communication to which they were accustomed. In addition, they were subjected to continuous exhortations (backed by threats of severe punishment) to confess their crimes and adopt new attitudes, and were constantly humiliated in order to discredit their old sense of identity.

The isolation of the target from his normal social and ideological supports reached its height in the case of Western civilians who were placed into group cells with a number of Chinese prisoners who had already confessed and were committed to reforming themselves and their lone Western cell mate. In the prisoner of war camps such extreme social isolation could not be produced, but its counterpart was created by the fomenting of mutual mistrust among the prisoners, by cutting off any supportive mail from home, and by systematically disorganizing the formal and informal social structure of the POW camp (by segregation of officers and non-commissioned officers from the remainder of the group, by the systematic removal of informal leaders or key personalities, and by the prohibition of any group activity not in line with the indoctrination program) (Schein, 1960, 1961).

The Chinese did not hesitate to use physical brutality and threats of death and/or permanent non-repatriation to enforce the view that only by collaboration and attitude change could the prisoner hope to survive physically and psychologically. In the case of the civilians in group cells, an additional and greater stress was represented by the social pressure of the cell mates who would harangue, insult, revile, humiliate, and plead with the resistant Westerner twenty-four hours a day for weeks or months on end, exhorting him to admit his guilt, confess his crimes, reform, and adopt Communist values. This combination of physical and social pressures is perhaps a prototype of the use of coercion in the service of unfreezing a target individual in attitude areas to which he is strongly committed.

A somewhat milder, though structurally similar, process can be observed in the training of a nun (Hulme, 1956). The novice enters the convent voluntarily and is presumably ready to change, but the kind of change which must be accomplished encounters strong psychological resistances because, again, it involves deeply held attitudes and habits. Thus the novice must learn to be completely unselfish and, in fact, selfless; she must adapt to a completely communal life; she must give up any source of authority except the absolute authority of God and of those senior to her in the convent; and she must learn to curb her sexual and aggressive impulses. How does the routine of the convent facilitate unfreezing? Again a key element is the removal of the novice from her accustomed routines, sources of confirmation, social supports, and old relationships. She is physically isolated from the outside world, surrounded by others who are undergoing the same training as she, subjected to a highly demanding and fatiguing physical regimen, constantly exhorted toward her new role and punished for any evidence of old behaviors and attitudes, and subjected to a whole range of social pressures ranging from mild disapproval to total humiliation for any failure.

Not only is the novice cut off from her old social identity, but her entry into the convent separates her from many aspects of her physical identity. She is deprived of all means of being beautiful or even feminine; her hair is cut off and she is given institutional garb which emphasizes formlessness and sameness; she loses her old name and chronological age in favor of a new name and age corresponding to length of time in the convent; her living quarters and daily routine emphasize an absolute minimum of physical comfort and signify a total devaluation of anything related to the body. At the same time the threat associated with change is minimized by the tremendous support which the convent offers for change and by the fact that everyone else either already exhibits the appropriate attitudes or is in the process of learning them.

If we look at the process by which a pledge comes to be a full-fledged member of a fraternity, we find in this situation also a set of pressures to give up old associations and habits, a devaluation of the old self by humiliations ranging from menial, senseless jobs to paddling and hazing, a removal of threat through sharing of training, and support for good performance in the pledge role. The evangelist seeking to convert those who come to hear him attempts to unfreeze his audience by stimulating guilt and by devaluating their former selves as sinful and unworthy. The teacher wishing to induce motivation to learn sometimes points out the deficiencies in the student's knowledge and hopes at the same time to induce some guilt for having those deficiencies.

Some of the elements which all unfreezing situations have in common are the following: (1) the physical removal of the influence target from his accustomed routines, sources of information, and social relationships; (2) the undermining and destruction of

all social supports; (3) demeaning and humiliating experience to help the target see his old self as unworthy and thus to become motivated to change; (4) the consistent linking of reward with willingness to change and of punishment with unwillingness to change.

Changing. Once the target has become motivated to change, the actual influence is most likely to occur by one of two processes. The target finds one or more models in his social environment and learns new attitudes by identifying with them and trying to become like them; or the target confronts new situations with an experimental attitude and develops for himself attitudes which are appropriate to the situation and which remove whatever problem he faces. These two processes - identification and internalization - probably tend to occur together in most concrete situations, but it is worthwhile, for analytical purposes, to keep them separate.*

The student or prisoner of the Chinese Communists took his basic step toward acquiring Communist attitudes when he began to identify with his more advanced fellow student or prisoner. In the group cell it was the discovery by the Western prisoner that his Chinese cell mates were humans like himself, were rational, and yet completely believed in their own and his guilt, which forced him to re-examine his own premises and bases of judgment and led him the first step down the path of acquiring the Communist point of view. In other words, he began to identify with his cell mates and to acquire their point of view as the only solution to getting out of prison and reducing the pressure on him. The environment was, of course, saturated with the Communist point of view, but it is significant that such saturation by itself was not sufficient to induce genuine attitude change. The prisoner kept in isolation and bombarded with propaganda was less likely to acquire Communist attitudes than the one placed into a group cell with more reformed prisoners. Having a personal model was apparently crucial.

In the convent the situation is essentially comparable except that the novice is initially much more disposed toward identifying with older nuns and has a model of appropriate behavior around her all the time in the actions of the others. It is interesting to note also that some nuns are singled out as particularly qualified models and given the appropriate name of "the living rule." It is also a common institution in initiation or indoctrination procedures to attach to the target individual someone who is labelled a "buddy" or "big brother," whose responsibility it is to teach the novice "the ropes" and to communicate the kinds of attitudes expected of him.

In most kinds of training and teaching situations, and even in the sales relationship, it is an acknowledged fact that the process is facilitated greatly if the target can identify with the influence

---

* Both are facilitated greatly if the influence agent saturates the environment with the new message or attitude to be learned.

agent. Such identification is facilitated if the social distance and rank difference between agent and target are not too great. The influence agent has to be close enough to the target to be seen as similar to the target, yet must be himself committed to the attitudes he is trying to inculcate. Thus, in the case of the Chinese Communist group cell, the cell mates could be perceived as sharing a common situation with the Western prisoner and this perception facilitated his identification with them. In most buddy systems, the buddy is someone who has himself gone through the training program in the recent past. If the target is likely to mistrust the influence attempts of the organization, as might be the case in a management-sponsored training program for labor or in a therapy program for delinquents in a reformatory, it is even more important that the influence agent be perceived as similar to the target. Otherwise he is dismissed as a "company man" or one who has already sold out, and hence is seen as someone whose message or example is not to be taken seriously.

Internalization, the discovery of attitudes which are the target's own solutions to his perceived dilemmas, can occur at the same time as identification. The individual can use the example of others to guide him in solving his own problems without necessarily identifying with them to the point of complete imitation. His choice of attitude remains ultimately his own in terms of what works for him, given the situation in which he finds himself. Internalization is only possible in an organizational context in which, from the organization's point of view, a number of different kinds of attitudes will be tolerated. If there is a "party line," a company philosophy, or a given way in which people have to feel about things in order to get along, it is hardly an efficient procedure to let trainees discover their own solutions. Manipulating the situation in such a way as to make the official solution the only one which is acceptable can, of course, be attempted, but the hazards of creating real resentment and alienation on the part of the individual when he discovers he really had no choice may outweigh the presumed advantages of letting him think he had a choice.

In the case of the Chinese Communists, the convent, the revival meeting, the fraternity, or the institutional training program, we are dealing with situations where the attitudes to be learned are clearly specified. In this kind of situation, internalization will not occur unless the attitudes to be learned happen to fit uniquely the kind of personal problem the individual has in the situation. For example, a few prisoners of the Communists reacted to the tremendous unfreezing pressures with genuine guilt when they discovered they held certain prejudices and attitudes (e.g. when they realized that they had looked down on lower class Chinese in spite of their manifest acceptance of them). These prisoners were then able to internalize certain portions of the total complex of Communist attitudes, particularly those dealing with unselfishness and working for the greater good of others. The attitudes which the

institution demanded of them also solved a personal problem of long standing for them. In the case of the nun, one might hypothesize that internalization of the convent's attitudes will occur to the extent that asceticism offers a genuine solution to the incumbent's personal conflicts.

Internalization is a more common outcome in those influence settings where the direction of change is left more to the individual. The influence which occurs in programs like Alcoholics Anonymous, in psychotherapy or counseling for hospitalized or incarcerated populations, in religious retreats, in human relations training of the kind pursued by the National Training Laboratories (1953), and in certain kinds of progressive education programs is more likely to occur through internalization or, at least, to lead ultimately to more internalization.

Refreezing. Refreezing refers to the process by which the newly acquired attitude comes to be integrated into the target's personality and ongoing relationships. If the new attitude has been internalized while being learned, this has automatically facilitated refreezing because it has been fitted naturally into the individual's personality. If it has been learned through identification, it will persist only so long as the target's relationship with the original influence model persists unless new surrogate models are found or social support and reinforcement is obtained for expressions of the new attitude.*

In the case of the convent such support comes from a whole set of expectations which others have of how the nun should behave, from clearly specified role prescriptions, and from rituals. In the case of individuals influenced by the Chinese Communists, if they remained in Communist China they received constant support for their new attitudes from superiors and peers; if they returned to the West, the permanence of their attitude change depended on the degree of support they actually received from friends and relations back home, or from groups which they sought out in an attempt to get support. If their friends and relatives did not support Communist attitudes, the repatriates were influenced once again toward their original attitudes or toward some new integration of both sets.

The importance of social support for new attitudes was demonstrated dramatically in the recent Billy Graham crusade in New York City. An informal survey of individuals who came forward when Graham called for converts indicated that only those individuals who were subsequently integrated into local churches maintained their faith. Similar kinds of findings have been repeatedly

---

* In either case the change may be essentially permanent, in that a relationship to a model or surrogate can last indefinitely. It is important to distinguish the two processes, however, because if one were to try to change the attitude, different strategies would be used depending upon how the attitude had been learned.

noted with respect to human relations training in industry. Changes which may occur during the training program do not last unless there is some social support for the new attitudes in the "back home" situation.

The kind of model which has been discussed above might best be described by the term "coercive persuasion." The influence of an organization on an individual is coercive in the sense that he is usually forced into situations which are likely to unfreeze him, in which there are many overt and covert pressures to recognize in himself a need for change, and in which the supports for his old attitudes are in varying degrees coercively removed. It is coercive also to the degree that the new attitudes to be learned are relatively rigidly prescribed. The individual either learns them or leaves the organization (if he can). At the same time, the actual process by which new attitudes are learned can best be described as persuasion. In effect, the individual is forced into a situation in which he is likely to be influenced. The organization can be highly coercive in unfreezing its potential influence targets, yet be quite open about the direction of attitude change it will tolerate. In those cases where the direction of change is itself coerced (as contrasted with letting it occur through identification or internalization), it is highly unlikely that anything is accomplished other than surface behavioral change in the target. And such surface change will be abandoned the moment the coercive force of the change agent is lessened. If behavioral changes are coerced at the same time as other unfreezing operations are undertaken, actual influence can be facilitated if the individual finds himself having to learn attitudes to justify the kinds of behavior he has been forced to exhibit. The salesman may not have an attitude of cynicism toward his customers initially. If, however, he is forced by his boss to behave as if he felt cynical, he might develop real cynicism as a way of justifying his actual behavior.

MANAGEMENT DEVELOPMENT: IS IT COERCIVE PERSUASION?

Do the notions of coercive persuasion developed above fit the management development situation? Does the extent to which they do or do not fit such a model illuminate for us some of the implications of specific management development practices?

Unfreezing. It is reasonable to assume that the majority of managers who are being "developed" are not ready or able to change in the manner in which their organization might desire and therefore must be unfrozen before they can be influenced. They may be eager to change at a conscious motivation level, yet still be psychologically unprepared to give up certain attitudes and values in favor of untried, threatening new ones. I cannot support this assumption empirically, but the likelihood of its being valid is high because of a related fact which is empirically supportable. Most managers do not participate heavily in decisions which affect their careers, nor do they have a large voice in the kind of self-develop-

ment in which they wish to participate. Rather, it is the man's superior or a staff specialist in career development who makes the key decisions concerning his career (Alfred, 1960). If the individual manager is not trained from the outset to take responsibility for his own career and given a heavy voice in diagnosing his own needs for a change, it is unlikely that he will readily be able to appreciate someone else's diagnosis. It may be unclear to him what basically is wanted of him or, worse, the ambiguity of the demands put upon him combined with his own inability to control his career development is likely to arouse anxiety and insecurity which would cause even greater resistance to genuine self-assessment and attitude change.* He becomes preoccupied with promotion in the abstract and attempts to acquire at a surface level the traits which he thinks are necessary for advancement.

If the decisions made by the organization do not seem valid to the manager, or if the unfreezing process turns out to be quite painful to him, to what extent can he leave the situation? His future career, his financial security, and his social status within the business community all stand to suffer if he resists the decisions made for him. Perhaps the most coercive feature is simply the psychological pressure that what he is being asked to do is "for his own ultimate welfare." Elementary loyalty to his organization and to his managerial role demands that he accept with good grace whatever happens to him in the name of his own career development. In this sense, then, I believe that the business organization has coercive forces at its disposal which are used by it in a manner comparable to the uses made by other organizations.

Given the assumption that the manager who is to be developed needs to be unfrozen, and given that the organization has available coercive power to accomplish such unfreezing, what mechanisms does it actually use to unfreeze potential influence targets?

The essential elements to unfreezing are the removal of supports for the old attitudes, the saturation of the environment with the new attitudes to be acquired, a minimizing of threat, and a maximizing of support for any change in the right direction. In terms of this model it becomes immediately apparent that training programs or other activities which are conducted in the organization at the place of work for a certain number of hours per day or week are far less likely to unfreeze and subsequently influence the participant than those programs which remove him for varying lengths of time from his regular work situation and normal social relationships.

---

* An even greater hazard, of course, is that the organization communicates to the manager that he is not expected to take responsibility for his own career at the same time that it is trying to teach him how to be able to take responsibility for important decisions!

Are appraisal interviews, used periodically to communicate to the manager his strengths, weaknesses and areas for improvement, likely to unfreeze him? Probably not, because as long as the individual is caught up in his regular routine and is responding, probably quite unconsciously, to a whole set of expectations which others have about his behavior and attitudes, it is virtually impossible for him to hear, at a psychological level, what his deficiencies or areas needing change are. Even if he can appreciate what is being communicated to him at an intellectual level, it is unlikely that he can emotionally accept the need for change, and even if he can accept it emotionally, it is unlikely that he can produce change in himself in an environment which supports all of his old ways of functioning. This statement does not mean that the man's coworkers necessarily approve of the way he is operating or like the attitudes which he is exhibiting. They may want to see him change, but their very expectations concerning how he normally behaves operate as a constraint on him which makes attitude change difficult in that setting.

On the other hand, there are a variety of training activities which are used in management development which approximate more closely the conditions necessary for effective unfreezing. These would include programs offered at special training centers such as those maintained by IBM on Long Island and General Electric at Crotonville, N. Y.; university-sponsored courses in management, liberal arts, and/or the social sciences; and especially, workshops or laboratories in human relations such as those conducted at Arden House, N. Y., by the National Training Laboratories. Programs such as these remove the participant for some length of time from his normal routine, his regular job, and his social relationships (including his family in most cases), thus providing a kind of moratorium during which he can take stock of himself and determine where he is going and where he wants to go.

The almost total isolation from the pressures of daily life in the business world which a mountain chateau such as Arden House provides for a two-week period is supplemented by other unfreezing forces. The de-emphasis on the kind of job or title the participant holds in his company and the informal dress remove some of the symbolic or status supports upon which we all rely. Sharing a room and bath facilities with a roommate requires more than the accustomed exposure of private spheres of life to others. The total involvement of the participant in the laboratory program leaves little room for reflection about the back home situation. The climate of the laboratory communicates tremendous support for any efforts at self-examination and attempts as much as possible to reduce the threats inherent in change by emphasizing the value of experimentation, the low cost and risk of trying a new response in the protected environment of the lab, and the high gains to be derived from finding new behavior patterns and attitudes which might improve back home performance. The content of the material pre-

sented in lectures and the kind of learning model which is used in the workshop facilitates self-examination, self-diagnosis based on usable feedback from other participants, and rational planning for change.*

The practice of rotating a manager from one kind of assignment to another over a period of years can have some of the same unfreezing effects and thus facilitate attitude change. Certainly his physical move from one setting to another removes many of the supports to his old attitudes, and in his new job the manager will have an opportunity to try new behaviors and become exposed to new attitudes. The practice of providing a moratorium in the form of a training program prior to assuming a new job would appear to maximize the gains from each approach, in that unfreezing would be maximally facilitated and change would most probably be lasting if the person did not go back to a situation in which his co-workers, superiors, and subordinates had stable expectations of how he should behave.

Another example of how unfreezing can be facilitated in the organizational context is the practice of temporarily reducing the formal rank and responsibilities of the manager by making him a trainee in a special program, or an apprentice on a special project, or an assistant to a high ranking member of the company. Such temporary lowering of formal rank can reduce the anxiety associated with changing and at the same time serves officially to destroy the old status and identity of the individual because he could not ordinarily return to his old position once he had accepted the path offered by the training program. He would have to move either up or out of the organization to maintain his sense of self-esteem. Of course, if such a training program is perceived by the trainee as an indication of his failing rather than a step toward a higher position, his anxiety about himself would be too high to facilitate effective change on his part. In all of the illustrations of organizational influence we have presented above, change was defined as being a means of gaining status--acceptance into Communist society, status as a nun or a fraternity brother, salvation, etc. If participants come to training programs believing they are being punished, they typically do not learn much.

The above discussion is intended to highlight the fact that some management development practices do facilitate the unfreezing of the influence target, but that such unfreezing is by no means automatic. Where programs fail, therefore, one of the first questions we must ask is whether they failed because they did not provide adequate conditions for unfreezing.

---

* Although, as I will point out later, such effective unfreezing may lead to change which is not supported or considered desirable by the "back home" organization.

Changing. Turning now to the problem of the mechanisms by which changes actually occur, we must confront the question of whether the organization has relatively rigid prescribed goals concerning the direction of attitude change it expects of the young manager, or whether it is concerned with growth in the sense of providing increasing opportunities for the young manager to learn the attitudes appropriate to ever more challenging situations. It is undoubtedly true that most programs would claim growth as their goal, but the degree to which they accomplish it can only be assessed from an examination of their actual practice.

Basically the question is whether the organization influences attitudes primarily through the mechanism of identification or the mechanism of internalization. If the development programs stimulate psychological relationships between the influence target and a member of the organization who has the desired attitudes, they are thereby facilitating influence by identification but, at the same time, are limiting the alternatives available to the target and possibly the permanence of the change achieved. If they emphasize that the target must develop his own solutions to ever more demanding problems, they are risking that the attitudes learned will be incompatible with other parts of the organization's value system but are producing more permanent change because the solutions found are internalized. From the organization's point of view, therefore, it is crucial to know what kind of influence it is exerting and to assess the results of such influence in terms of the basic goals which the organization may have. If new approaches and new attitudes toward management problems are desired, for example, it is crucial that the conditions for internalization be created. If rapid learning of a given set of attitudes is desired, it is equally crucial that the conditions for identification with the right kind of models be created.

One obvious implication of this distinction is that programs conducted within the organization's orbit by its own influence agents are much more likely to facilitate identification and thereby the transmission of the "party line" or organization philosophy. On the other hand, programs like those conducted at universities or by the National Training Laboratories place much more emphasis on the finding of solutions by participants which fit their own particular needs and problems. The emphasis in the human relations courses is on "learning how to learn" from the participant's own interpersonal experiences and how to harness his emotional life and intellectual capacities to the accomplishment of his goals, rather than on specific principles of human relations. The nearest thing to an attitude which the laboratory staff, acting as influence agents, does care to communicate is an attitude of inquiry and experimentation, and to this end the learning of skills of observation, analysis, and diagnosis of interpersonal situations is given strong emphasis. The training group, which is the acknowledged core of the laboratory approach, provides its own unfreezing forces by

being unstructured as to the content of discussion. But it is strongly committed to a method of learning by analysis of the member's own experiences in the group, which facilitates the discovery of the value of an attitude of inquiry and experimentation.

Mutual identification of the members of the group with each other and member identifications with the staff play some role in the acquisition of this attitude, but the basic power of the method is that the attitude of inquiry and experimentation works in the sense of providing for people valuable new insights about themselves, groups, and organizations. To the extent that it works and solves key problems for the participants, it is internalized and carried back into the home situation. To the extent that it is learned because participants wish to emulate a respected fellow member or staff member, it lasts only so long as the relationship with the model itself, or a surrogate of it, lasts (which may, of course, be a very long time).

The university program in management or liberal arts is more difficult to categorize in terms of an influence model, because within the program there are usually opportunities both for identification (e.g. with inspiring teachers) and internalization. It is a safe guess in either case, however, that the attitudes learned are likely to be in varying degrees out of phase with any given company's philosophy unless the company has learned from previous experience with a given course that the students are taught a point of view consistent with its own philosophy. Of course, universities, as much as laboratories, emphasize the value of a spirit of inquiry and, to the extent that they are successful in teaching this attitude, will be creating potential dissidents or innovators, depending on how the home company views the result.

Apprenticeships, special jobs in the role of "assistant to" somebody, job rotation, junior management boards, and so on stand in sharp contrast to the above methods in the degree to which they facilitate, indeed almost demand, that the young manager learn by watching those who are senior or more competent. It is probably not prescribed that in the process of acquiring knowledge and skills through the example of others he should also acquire their attitudes, but the probability that this will happen is very high if the trainee develops any degree of respect and liking for his teacher and/or supervisor. It makes little difference whether the teacher, coach, or supervisor intends to influence the attitudes of his trainee or not. If a good emotional relationship develops between them, it will facilitate the learning of knowledge and skills, and will, at the same time, result in some degree of attitude change. Consequently, such methods do not maximize the probability of new approaches being invented to management problems, nor do they really by themselves facilitate the growth of the manager in the sense of providing opportunities for him to develop solutions which fit his own needs best.

Job rotation, on the other hand, can facilitate growth and innovation provided it is managed in such a way as to insure the exposure of the trainee to a broad range of points of view as he moves from assignment to assignment. The practice of shifting the developing manager geographically as well as functionally both facilitates unfreezing and increases the likelihood of his being exposed to new attitudes. This same practice can, of course, be merely a convenient way of indoctrinating the individual by sending him on an assignment, for example, "in order to acquire the sales point of view from Jim down in New York," where higher management knows perfectly well what sort of a view Jim will communicate to his subordinates.

Refreezing. Finally, a few words are in order about the problem of refreezing. Under what conditions will changed attitudes remain stable, and how do existing practices aid or hinder such stabilization? Our illustrations from the non-industrial setting highlighted the importance of social support for any attitudes which were learned through identification. Even the kind of training emphasized in the National Training Laboratories programs, which tends to be more internalized, does not produce stable attitude change unless others in the organization, especially superiors, peers, and subordinates, have undergone similar changes and give each other stimulation and support, because lack of support acts as a new unfreezing force producing new influence (possibly in the direction of the original attitudes).

If the young manager has been influenced primarily in the direction of what is already the company philosophy, he will, of course, obtain strong support and will have little difficulty maintaining his new attitudes. If, on the other hand, management development is supposed to lead to personal growth and organizational innovation, the organization must recognize the reality that new attitudes cannot be carried by isolated individuals. The lament that we no longer have strong individualists who are willing to try something new is a fallacy based on an incorrect diagnosis. Strong individuals have always gained a certain amount of their strength from the support of others, hence the organizational problem is how to create conditions which make possible the nurturing of new ideas, attitudes, and approaches. If organizations seem to lack innovators, it may be that the climate of the organization and its methods of management development do not foster innovation, not that its human resources are inadequate.

An organizational climate in which new attitudes which differ from company philosophy can nevertheless be maintained cannot be achieved merely by an intellectual or even emotional commitment on the part of higher-ranking managers to tolerance of new ideas and attitudes. Genuine support can come only from others who have themselves been influenced, which argues strongly that at least several members of a given department must be given the same training before such training can be expected to have effect.

If the superior of the people involved can participate in it as well, this strengthens the group that much more, but it would not follow from my line of reasoning that this is a necessary condition. Only some support is needed, and this support can come as well from peers and subordinates.

From this point of view, the practice of sending more than one manager to any given program at a university or human relations workshop is very sound. The National Training Laboratories have emphasized from the beginning the desirability of having organizations send teams. Some organizations like Esso Standard have created their own laboratories for the training of the entire management complement of a given refinery, and all indications are that such a practice maximizes the possibility not only of the personal growth of the managers, but of the creative growth of the organization as a whole.

CONCLUSION

In the above discussion I have deliberately focused on a model of influence which emphasizes procedure rather than content, interpersonal relations rather than mass media, and attitudes and values rather than knowledge and skills. By placing management development into a context of institutional influence procedures which also include Chinese Communist thought reform, the training of a nun, and other more drastic forms of coercive persuasion, I have tried to highlight aspects of management development which have remained implicit yet which need to be understood. I believe that some aspects of management development are a mild form of coercive persuasion, but I do not believe that coercive persuasion is either morally bad in any a priori sense nor inefficient. If we are to develop a sound theory of career development which is capable of including not only many of the formal procedures discussed in this paper, but the multitudes of informal practices, some of which are more and some of which are less coercive than those discussed, we need to suspend moral judgments for the time being and evaluate influence models solely in terms of their capacity to make sense of the data and to make meaningful predictions.

REFERENCES

Alfred, T. M. Personal communication, 1960.

Hulme, K. The nun's story. Boston, Little, Brown, 1957.

Kelman, H. C. Compliance, identification, and internalization: three processes of attitude change. Conflict Resolution, 1958, 2, 51-60.

Lewin, K. Frontiers in group dynamics: concept, method and reality in social science. Hum. Relat., 1947, I, 5-42.

McGregor, D. The human side of enterprise. New York, McGraw-Hill, 1960.

Merton, R. K., Reader, G. G., and Kendall, Patricia L. The student-physician. Cambridge, Mass., Harvard University Press, 1957.

National Training Laboratory in Group Development: Explorations in human relations training: an assessment of experience, 1947-1953. Washington, D. C., National Education Association, 1953.

Schein, E. H. Brainwashing. Cambridge, Mass., Center for International Studies, M.I.T., 1961.

Schein, E. H. Interpersonal communication, group solidarity, and social influence. Sociometry, 1960, 23, 148-161.